Ms Fanny Sarrazin
Research Associate, Water and Environmental Engineering
Department of Civil Engineering, University of Bristol
Queen's Building, University Walk, Bristol BS81TR, UK
fanny.sarrazin@bristol.ac.uk

Dear Dr Roche,

Please find enclosed a revised version of our manuscript "V2Karst V1.0: A parsimonious large-scale integrated vegetation-recharge model to simulate the impact of climate and land cover change in karst regions" in which we address the Reviewers' comments. The Reviewers' detailed comments helped us to significantly improve our manuscript, although we have found it difficult to address some of Reviewer #2's specific comments, because he/she did not provide any reference to support his/her statements. We attached a detailed reply to the Reviewers' comments and a marked-up version of the manuscript that highlights in red the changes we made. Since we substantially revised our manuscript, we did not keep track of minor editorial changes to facilitate readability of the marked-up manuscript. Specifically, we wish to report some key changes:

- To address the concern of Reviewer #2 regarding the suitability of using a daily simulation time step to run V2Karst, we performed additional simulations at hourly time step (reported in our Supplementary material). Our results suggest that there are no significant differences between V2Karst predictions for daily and hourly time steps for our study purpose, i.e. simulation of monthly and annual recharge.
- We clarified our rationale for development of our parsimonious large-scale hydrological model (Sect. 2.1). We followed an approach to large-scale hydrological modelling that has been widely used for climate change impact studies. We aimed to limit model complexity given the limited amount of information available at large-scales to constrain model simulations.
- Following the advice of Reviewer #1, we made an effort to shorten the manuscript and improve its readability, while also addressing the Reviewers' comments.
- We rewrote the discussion section to better convey the meaning and implications of our sensitivity analyses results and to discuss large-scale applications of V2Karst (Sect. 6).

Again, we wish to thank you and the Reviewers' for time and effort taken to review our manuscript and for any further comment and suggestion.

Sincerely,

Fanny Sarrazin

F Sarrazin

Response to Reviewer #1

COMMENT 1: *The manuscript presents a modified version of the large-scale karst recharge model VarKarst. The here presented model (V2Karst V1.0) replaces the simplified evapotranspiration (ET) component (empirical Priestley-Taylor equation) by the physical based Penman-Monteith equation (for potential evapotranspiration). The authors also include a separate calculation of the different evaporation processes in order to use the model for climate and land cover change impact studies. The model extension increases the number of parameters. The general functioning as well as the influence of the new parameters are tested by applying the new model to four study sides, different in climate and vegetation. The manuscript is a novel extension of previous work published by the research group. The conceptual description and the numerical adaptation of the processes are sound. The results of the model application on the four test sides prove the general functioning of the new model. However, the manuscript also has weak points, which are mainly related to the presentation of the method and the results. The manuscript can easily be shortened by 10-20% without losing important information. The presentation of the results needs to be improved, especially since it is difficult to distinguish between observed values and modeled results. My detailed comments are listed below.*

REPLY 1: We wish to thank the Reviewer for appreciating the novelty and soundness of our work and for the detailed comments, especially regarding the karst aspect of our work. **Line numbers in our replies refer to the marked-up version of our revised manuscript** (included in this file). In the marked-up version of the manuscript we highlighted in red the insertion and deletion we made. We took into account the Reviewers' concerns regarding the length of the manuscript and the presentation of the methods and the results as explained in the following.

1. Shortening the manuscript

We realise that the manuscript was rather long in parts. We made an effort to shorten our revised version and to improve its readability and in particular:

- We followed the suggestions of the Reviewer (COMMENT 6). We revised the introductions of the sections to make them more informative (Sect. 4 p19 L10-19) and deleted them where appropriate (Sect. 2 p5 L29-30, in Sect. 2 p8 L3-5, Sect. 5 p23 L21-23). We deleted wordy expressions such as 'Regarding the data processing' p18 L9, 'using numerical models' p22 L6 and 'the sensitivity indices analysed in this study' that we replaced by 'we analyse' p21 L15. We rewrote wordy sentences such as p26 L27-28.
- In Sect. 1, we rewrote the paragraph that describes the objectives of the study p4 L31-p5 L12 more concisely and deleted unnecessary repetitions. We also shorten the sentence p5 L13 by deleting 'that help us to overcome the previous limitation'.
- In Sect. 2.1 we removed the expression p6 L5-6 'and uncertainty in large-scale weather forcing variables […]', which is not further discussed in our manuscript. We simplified the paragraph p6 L31-p7 L3. We removed the sentence p7 L13-14 which provides unnecessary details on a previous karst model.
- In Sect. 3.2, we deleted the reference to the 'residual estimate' of ET p18 L29-p19 L8 and we now only discuss the 'bowen ratio estimate' p18 L15-18 that we actually used in our analyses. Further analyses of the uncertainty in ET observations using the 'residual estimate' are reported in Sect. S4 of our Supplementary material as indicated p18 L19-21.
- In Sect. 5, we deleted more specifically the sentences p24 L24-25 and p27 L6-8 that pertained to Sect. 6 (discussion section). We moved the sentence p26 L15-16 into Sect. S7 of our Supplementary material, since it is a detailed comment on additional sensitivity analysis results that are reported in Sect. S7.

- We rewrote Sect. 6 more concisely, while also clarifying the meaning and implications of our sensitivity analyses and addressing other comments of the Reviewers (in particular COMMENT 3 and 5 below).

2. Presenting the methods and the results

We clarified the presentation of the methods and the results and specifically:

- We revised the introduction to the method section (Sect. 4) to make it more informative and to clarify the distinction between observed and simulated values (p19 L10-19). In brief, the parameter estimation approach (Sect. 4.1) uses measurements of weather variables to force the model and measurements of model output to define the soft rules. The global sensitivity analysis (Sect 4.2) uses measurements of weather variables to force the model, but no measurements of model output. Finally, the virtual experiments (Sect 4.3) do not use any measurements but synthetic data to force the model.
- We rewrote parts of Sect. 4.3 to clarify the set-up of the virtual experiments p23 L4-18.
- We prepared a new version of Fig. 4 in which we corrected some plotting issues (as mentioned in COMMENT 13), which may bring some clarity to the meaning of this figure.
- We simplified Fig. 5 which may have brought some confusion regarding the distinction between observed and simulated values. We now only report the bowen ratio estimate of observed ET of Eq. (9) p18 and we removed the lines that corresponded to the other estimates of observed ET (estimate without correction and with residual correction) that are now reported in Fig. S2 in our Supplementary material. We clarified in the caption of Fig. 5 and p24 L26-28 the variables that are observed and simulated.
- We also made numerous edits to the results section (Sect. 5) to improve its readability, for instance in the presentation of the results of the virtual experiments p28 L15-27.

MAIN COMMENTS

COMMENT 2: *The purpose of V2Karst V1.0 is to predict recharge in karst regions. The authors mention that "a large part of the groundwater recharge occurs as concentrated and fast flow in large apertures and the other part as diffuse and slow flow in the matrix (Hartmann and Baker, 2017)." Especially concentrated recharge, e.g. fast infiltration into sinkholes, can be considered as a short-term process and is entirely uncoupled from soil and/or vegetation properties (overland flow -> percolation). I assume that your model, calculating the water balance, underestimates the recharge in karst regions dominated by concentrated recharge. Do you think your model is able to equally represent both recharge processes?*

REPLY 2: We agree with the need for representing concentrated recharge processes. In karst systems, infiltration and recharge can be slow and diffuse in the matrix and fast and concentrated in large conduits or fissures that act as preferential flow pathways. Lateral flow at the surface and in the epikarst is an important mechanism that concentrates the infiltrating water into the preferential flow pathways (Jeannin and Grasso, 1997; Williams, 1983, 2008). In particular, the epikarst plays a role of temporary storage that can redistribute fast and concentrated recharge (Williams, 1983, 2008). Figure R1.c below provides a conceptual model of the soil and epikarst processes of a real karst system.

V2karst's representation of concentrated and diffuse infiltration and recharge, which is the same as in the VarKarst model (Hartmann et al., 2015), follows this conceptual model (Figure R1.a). V2karst represents the spatial variability of subsurface properties observed in karst systems by dividing each model simulation unit into a number of vertical compartments that have different soil and epikarst properties. Parameters for each model compartment (soil and epikarst storage capacities and epikarst outflow coefficient) are estimated using a distribution function. The daily water balance is explicitly evaluated for each model vertical compartment. Additionally, when a given compartment saturates (both soil and epikarst stores), its saturation excess generates a surface lateral flow to the next unsaturated compartments that have higher storage capacities and higher permeabilities. In this representation, surface and subsurface lateral flow are thus lumped together. Conceptually, in V2Karst, the direct contribution of precipitation to infiltration and recharge can be associated with diffuse infiltration and recharge, while the contribution of lateral flow can be associated with concentrated infiltration and recharge. Hence, the V2karst structure allows to account for the interplay between diffuse and concentrated infiltration and recharge processes.

The representation of karst processes in V2Karst and VarKarst is based on a previous karst model developed for applications at the local scale introduced in Hartmann et al. (2012). The structure of this previous models explicitly represents lateral flow both at the surface and in the epikarst, in agreement with understanding of the flow mechanisms in the epikarst (e.g. Williams, 1983, 2008). It was tested using hydrodynamic and hydrochemical information at stalactite drips in a karstic cave in Hartmann et al. (2012). However, simplifications have been introduced for applications at the large-scale, given the limited information available to constrain the additional parameters required in the previous karst model to represent lateral flow in the epikarst.

We are aware of the fact that, in V2Karst's and VarKarst's representation, concentrated recharge is not entirely uncoupled from the soil and vegetation properties as it is observed in real karst systems. However, a previous study by Hartmann et al. (2017) compared simulated recharge with VarKarst and independent estimates of recharge and showed that there was no systematic bias in the simulations (Figure R2 below). Moreover, the study by Hartmann et al. (2017) showed that recharge values simulated with VarKarst were significantly higher than recharge values simulated with models that do not include karst processes (Figure R2 below).

As also stated in REPLY 9, in the revised version of our manuscript, we briefly explained how diffuse and concentrated recharge is represented in VarKarst and V2karst in Sect. 2.2. p8 L12-17.

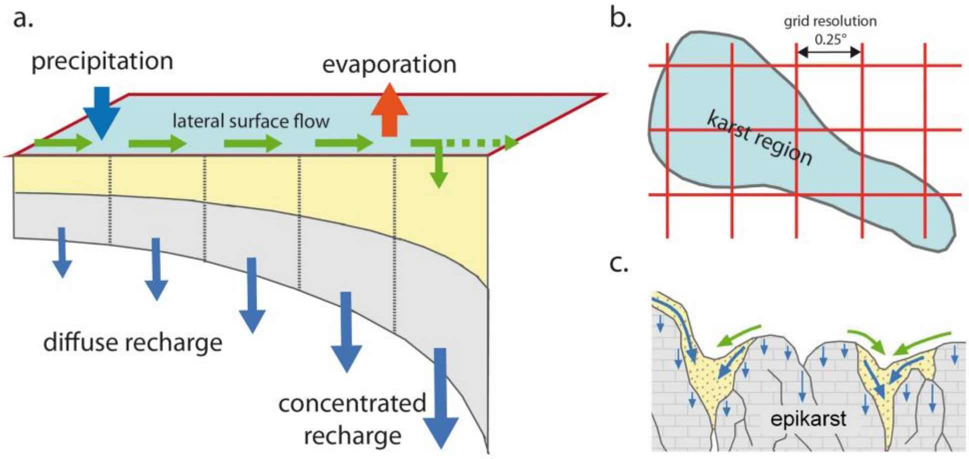


Figure R1. (a) Schematic description of the VarKarst model for one model grid cell including the soil (yellow) and epikarst storages (grey) and the simulated fluxes, (b) its gridded discretization over karst regions and (c) the subsurface heterogeneity that its structure represents for each grid cell.

Figure taken from Hartmann et al. (2015, Figure 1)

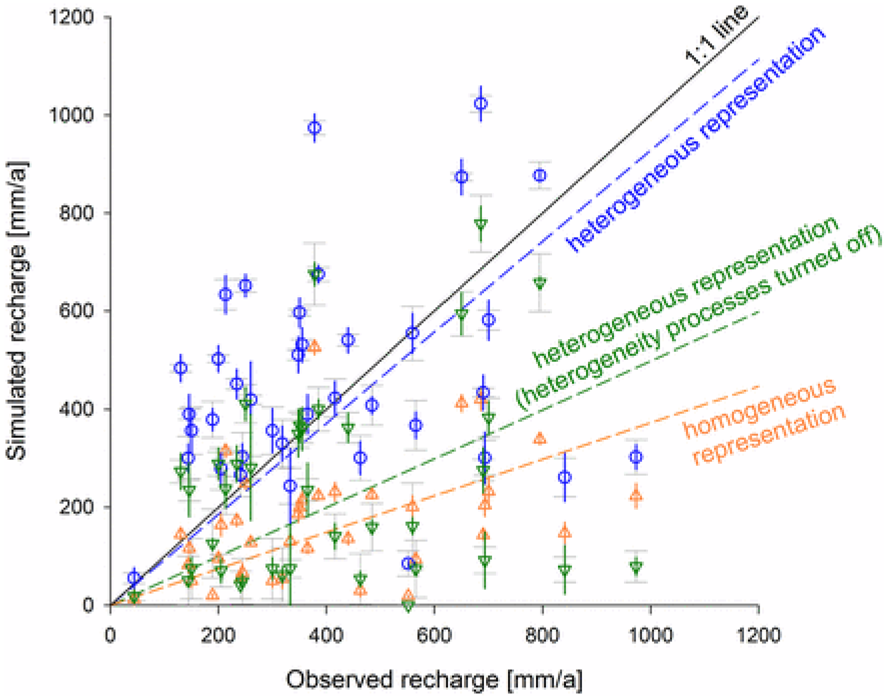


Figure R2. Comparison of simulated and observed recharge. **In blue are reported values simulated with the VarKarst model (heterogeneous representation)**, in yellow values simulated with the PCR-GLOBWB model (homogeneous representation, i.e. absence of karst processes in the model representation), and in green the Varkarst model with subsurface heterogeneity processes turned off. Whiskers indicate the simulation uncertainty (1 SD) for simulations with VarKarst. **No statistical difference (5% significance level) was found between simulations with Varkarst and the observations**. mm/a, millimeters per year.

Figure taken from Hartmann et al. (2017, Figure 2)

COMMENT 3: *I am aware of the fact that the manuscript is focused on the implementation and the testing of the new evapotranspiration component. Since soil layers in karst regions can be thin or even totally absence the authors should consider this fact in the interpretation of the results.*

REPLY 3: Large differences in soil depth can indeed be observed across different karst landscapes. To apply the VarKarst model over Europe and the Mediterranean, Hartmann et al. (2015) have identified four main karst landscapes with different soil depths and degrees of karstification based on climate and topography (Humid, Mountain, Mediterranean and Desert landscapes). Different values of the parameter V_{soi} (mean soil water capacity) have been applied in the different landscapes. In particular, very small values of V_{soi} have been used in arid areas (i.e. desert landscape in Hartmann et al. (2015)),where soils tend to be very thin.

We are also aware of the fact that soils may be absent in some karst areas (e.g. karren field in high mountain areas, see e.g. Hartmann et al. (2014a)), and that these areas may consequently produce very high recharge amounts. The model does not account for the fact that the soil may be absent and always includes a soil layer, although the soil layer can be very thin and therefore can have a limited impact on recharge. This assumption seems reasonable given the large extent of the simulation units for large-scale applications (0.25°x0.25° or 0.5°x0.5° cell in previous applications). In fact, we can assume that soil layers can always be found in such large simulation units. However, for model applications at high resolutions, we recognise the fact that an explicit consideration of bare rock regions should be included in the model.

We added a brief discussion of this point in Sect. 6.3 of our revised manuscript p33 L6-9.

COMMENT 4: *The manuscript lacks a description/characterization of the four karst regions (e.g. by describing dominant karst features or by the interpretation of spring hydrographs).*

> **REPLY 4:** In this study, we focus on groundwater recharge, which is a key component of the water balance as we clarified in our revised manuscript p5 L17-19. Recharge characterises the amount of renewable groundwater, and therefore the amount of groundwater available to human consumption and ecosystems (e.g. Scanlon et al., 2006; Döll and Fiedler, 2008; Wada et al., 2012). We do not model groundwater flow and storage nor spring discharge. Therefore, to test the model, we focus on datasets that can be related to the fluxes and states simulated by V2Karst, i.e. soil moisture and evapotranspiration as in Hartmann et al. (2015). No datasets providing time series of recharge are available. We do not use spring hydrographs, because spring discharge depends on groundwater routing and is not commensurate with groundwater recharge. Spring discharge measurements were used to extensively test a previous versions of the VarKarst model including groundwater storage and routing to the spring, at different locations in Europe and the Mediterranean (Hartmann et al., 2013a, 2013b, 2014b).
>
> In the manuscript, we briefly describe the sites in Sect. 3.1 p17. Table B1 describes the four sites in detail, and more specifically the soil depth and bedrock. We have realised that we have exchanged the name of the two French sites in Table B1 (Puéchabon is actually the French 2 site and Font-Blanche the French 1 site). We corrected this mistake in the revised version of the manuscript. We did not add more details on the sites in the main text to limit the length of the manuscript.

COMMENT 5: *In general, a differentiation between different karst systems and therefore the wide variety of*

> **REPLY 5**: A large variability in hydraulic properties and in recharge patterns can indeed be observed across karst systems (e.g. Klimchouk and Ford, 2000; Hartmann et al., 2014a). For applications over large-scale domain, Hartmann et al. (2015) identified four typical karst landscapes with different hydraulic properties, based on climate and topography (as mentioned in REPLY 3). This simplified classification of the simulation domain in four landscapes was introduced to enable large-scale applications of the model.
>
> In Sect. 6.3 (discussion of V2karst large-scale application), we better highlighted the fact that future large-scale applications of V2Karst will need to account for the variability observed across karst systems, for instance using a simplified classification as in Hartmann et al. (2015) p33 L1-4.

COMMENT 6: *As already mentioned, the current manuscript is too long and needs to be shortened:*
1) *(Almost) every section starts with a short introduction on the section. Most of them are redundant.*
2) *The authors use wordy descriptions instead of clear words for describing their work. Is a "virtual experiment with synthetic data to assess the sensitivity" (Page 1, Line 19) not simply a "sensitivity analysis"?*
3) *Discussion chapter: Consists of sentences/paragraphs, which can be defined as general knowledge (e.g. Page 26. Line 16; Page 26, Line 24) or which should be familiar by the readers of the journal (e.g. Page 28, Line 12).*

> **REPLY 6:** We thank the Reviewer for suggesting how to shorten the manuscript and we reply below to the three points raised by the Reviewer.

1) As already stated in REPLY 1, we revised the introductions of the sections to make them more informative (Sect. 4 p19 L10-19) and deleted them where appropriate (Sect. 2 p5 L29-30, in Sect. 2 p8 L3-5, Sect. 5 p23 L21-23).

2) The virtual experiments are indeed sensitivity analyses. However, an important aspect is that we used synthetic data, which has been done only in few studies (e.g. the studies reported p22 L12-15). The reason for using synthetic data is that it allows to explore conditions beyond what was historically observed, and it is therefore useful to better understand potential impact of changes in climate. It also permits unequivocal attribution of changes in model output to changes in model inputs. Therefore, we think that it is important to mention that we used synthetic data in the abstract p1 L23. However, as discussed in REPLY 7, we made substantial changes to the abstract for clarification and readability, including the discussion of the virtual experiments.
Furthermore, as stated in REPLY 1, we deleted wordy expressions in the manuscript, such as 'Regarding the data processing' p18 L9, 'using numerical models' p22 L6 and 'the sensitivity indices analysed in this study' that we replaced by 'we analyse' p21 L15. We rewrote wordy sentences such as p26 L27-28.

3) We agree that the discussion section can be shortened and clarified and more specifically at the places indicated by the Reviewer. We rewrote Sect. 6 more concisely, while also clarifying the meaning and implications of our sensitivity analyses in Sect. 6.1 and 6.2 p29-32 and addressing other comments of the Reviewers (including COMMENT 3 and 5 above).
We clarified the discussion of the virtual experiments (Sect. 6.2 p31-32). Firstly, the virtual experiments confirm that the model behaves reasonably, since it shows sensitivities to both precipitation (overall amount and temporal distribution) and land cover, in agreement with previous studies and general understanding. Secondly, the virtual experiments allow to unequivocally and quantitatively characterise the relationship between simulated recharge and both precipitation (overall mean and temporal distribution) and land cover. Therefore, virtual experiments are a complementary approach to model applications using site-specific data to assess the impact of climate and land cover changes on simulated recharge.

SECONDARY COMMENTS

COMMENT 7: - *Page 1, Line 21: ". . . and they suggest that simulated recharge is sensitive to both precipitation (overall amount and temporal distribution) and land cover." Is this one of the main results of your work and is it really a new finding?*

REPLY 7: We agree that we need to reformulate this sentence. We refer to REPLY 6 for a clarification of the objectives of the virtual experiments.
We rewrote this part of the abstract (see p1 L23-27 in the revised version of the manuscript).

COMMENT 8:
- *Page 2, Line 30: The sentence is difficult to understand.*
- *Page 4, Line 17: Please, rephrase the long sentence (and consider deleting the first part of the sentence).*

REPLY 8: We deleted the fist sentence mentioned by the Reviewer p3 L29-32. The split in two the second sentence mentioned by the Reviewer (p6 L2-7). We also removed the expression p6 L5-6 'and uncertainty in large-scale weather forcing variables […]', which is not further discussed in our manuscript.

COMMENT 9: *Page 6, Line 12: Could you please add a bit more information on how diffuse and concentrated recharge is considered by the model.*

REPLY 9: We refer to REPLY 2 for a detailed explanation of the conceptualisation of recharge processes in V2Karst. In the revised version of our manuscript, we briefly explained how diffuse and concentrated recharge is represented in VarKarst and V2karst in Sect. 2.2. p8 L12-17.

COMMENT 10: *Page 12, Line 12/13/14: Please, consider using SI-Units.*

REPLY 10: We have adopted these units because they are typically used in the evapotranspiration literature (e.g. Allen et al., 1998; Shuttleworth, 1993, 2012).
The revised version of the manuscript includes the unit of net radiation (Rn [MJ m^{-2} d^{-1}]) p15 L10 which was missing.

COMMENT 11: *Page 13 Seasonality of vegetation: Are you using the same seasonality on every study site irrespective of the local climate and vegetation type?*

REPLY 11: Typically, in hydrological models, vegetation seasonality is represented using different schemes with different levels of complexity or is neglected completely (Table A3). We chose to implement a simple representation (piecewise linear function). In this way, we can test the impact of vegetation seasonality by assessing the sensitivity of recharge to the seasonality parameter LAImin. We applied the same function at all sites, but we used different values of the seasonality parameter LAImin as indicated in Table 3. The timings of the four phases reported in Eq. (16) are appropriate for study sites that are located in the northern hemisphere and that have natural vegetation.

We added a sentence p16 L18-19 to clarify the fact that the timings of the four phases of the seasonality model should be adapted to the application domain. We also added 'in this study' p16 L11 and L12 to clarify the fact that the choice of growing and dormant season is appropriate for application at the sites used in this study.

COMMENT 12:
- *Page 15, Line 4: Please consider splitting the sentence.*
- *Equation 17/18: Please remove the units from the equation and mention both parameters in the text, e.g. "1. Eact,bow [mmmonth1], a corrected value that assumes that latent heat (ð˙IR£ð ˇ ˙IRÿ [MJ.m ˇ −2.month−1]) and sensible heat (ð˙IRˇ z [MJ.m ˙ −2.month−1]) have similar errors (referred to as Bowen ratio estimate):*
- *Page 17, Line 29: Please, rephrase the sentence.*
- *Page 24, Line 4: Please, rephrase the sentence.*

REPLY 12: We have applied these changes in the revised manuscript:
- p18 L2-4
- p18 L18
- p21 L5-7
- p27 L29-p28 L2

COMMENT 13: *Figure 4: The Figure presents the results in a confusing way and some of the values exceed the constrained parameter ranges according to Table 3.*

REPLY 13: We realise that the plot we used – a parallel coordinate plot – is not familiar to all readers in earth sciences. A parallel coordinate plot is a two-dimensional plot that allows to visualise a

multidimensional space (here the space of the model parameters and outputs). In Figure 4, each line represents a combination of model parameters values (normalised) and the corresponding model output values (normalised). Parallel coordinate plots are increasingly used (e.g. Inselberg, 2009; Kasprzyk et al., 2013; Pianosi et al., 2017), and are implemented for instance in the Matlab Statistics and Machine Learning Toolbox (function "parallelcoords") and in the SAFE toolbox for sensitivity analysis we utilised in our paper (Pianosi et al., 2015).
Initially, we sampled the parameter space within wide ranges (Table 2), and we applied a priori information on parameter ranges (Table 3) only in rule 5. Therefore, prior to application of rule 5, parameter values can exceed the ranges of Table 3 (yellow, light blue, green and dark blue lines in Figure 4). Instead, posterior to application of rule 5, all parameter values should be within the ranges of Table 3 (red lined in Figure 4).

The Reviewer may be referring to the fact that some red lines slightly exceed the black vertical lines in Figure 4. This is a plotting issue and we corrected it in the revised version of the manuscript.

MINOR COMMENTS AND TYPOGRAPHICAL ERRORS

COMMENT 14: *Please, use a consistent citation style.*

REPLY 14: We corrected the style of the reference p6 L7.

COMMENT 15: *Please, use a consistent style for figure references. Two different versions exist: Fig. and Figure.*

REPLY 15: We replaced 'Figure' by 'Fig.' p6 L22, p8 L26, p17 L11 so that the referencing of the figures is consistent with guidelines of GMD available at (https://www.geoscientific-model-development.net/for_authors/manuscript_preparation.html): *The abbreviation "Fig." should be used when it appears in running text and should be followed by a number unless it comes at the beginning of a sentence, e.g.: "The results are depicted in Fig. 5. Figure 9 reveals that...".*

COMMENT 16:
- *Units -> Replace the dots by multiplication sign or even better delete them.*
- *Page 2, Line 2: . . .world. . .*
- *Page 2, Line 2: For instance, . . .*
- *Page 5, Line 4: . . . (Hartmann et al., 2015). This . . . (space missing)*
- *Page 5, Line 33: . . . to represent. . .*
- *Page 7, Line 13: . . . the following formulas . . .*
- *Page 12, Line 13: . . . is the psychrometric constant, . . .*
- *Page 15, Line 18: . . . data processing are reported . . .*
- *Page 20, Line 18: red lines -> the a priori information are indicated by black lines in Figure 4!*
- *Page 29, Line 14: We, therefore, . . .*
- *Figure 5, Line 4: . . . percentage of Eact . . .*
- *Figure 6: Line 5: Blue . . .*
- *Figure 9: Line 4: remove the open bracket*

REPLY 16: The revised version of the manuscript will include these corrections. We chose to add 'More importantly' p2 L2.

References to support our reply to Reviewer #1

Allen, R. G., Pereira, L. S., Raes, D. and Smith, M.: Crop evapotranspiration: Guidelines for computing crop requirements, FAO Irrigation and Drainage Paper 56, Food and Agriculture Organization (FAO), Rome, Italy., 1998.

Döll, P. and Fiedler, K.: Global-scale modeling of groundwater recharge, Hydrol. Earth Syst. Sci., 12(3), 863–885, doi:10.5194/hess-12-863-2008, 2008.

Hartmann, A., Lange, J., Weiler, M., Arbel, Y. and Greenbaum, N.: A new approach to model the spatial and temporal variability of recharge to karst aquifers, Hydrol. Earth Syst. Sci., 16(7), 2219–2231, doi:10.5194/hess-16-2219-2012, 2012.

Hartmann, A., Weiler, M., Wagener, T., Lange, J., Kralik, M., Humer, F., Mizyed, N., Rimmer, A., Barberá, J. A., Andreo, B., Butscher, C. and Huggenberger, P.: Process-based karst modelling to relate hydrodynamic and hydrochemical characteristics to system properties, Hydrol. Earth Syst. Sci., 17(8), 3505–3521, doi:10.5194/hess-17-3305-2013, 2013a.

Hartmann, A., Barberá, J. A., Lange, J., Andreo, B. and Weiler, M.: Progress in the hydrologic simulation of time variant recharge areas of karst systems - Exemplified at a karst spring in Southern Spain, Adv. Water Resour., 54, 149–160, doi:10.1016/j.advwatres.2013.01.010, 2013b.

Hartmann, A., Goldscheider, N., Wagener, T., Lange, J. and Weiler, M.: Karst water resources in a changing world: Review of hydrological modeling approaches, Rev. Geophys., 52(3), 218–242, doi:10.1002/2013RG000443, 2014a.

Hartmann, A., Mudarra, M., Andreo, B., Marin, A., Wagener, T. and Lange, J.: Modelling spatiotemporal impacts of hydroclimatic extremes on groundwater recharge at a Mediterranean karst aquifer, Water Resour. Res., 50(8), 6507–6521, doi:10.1002/2014WR015685, 2014b.

Hartmann, A., Gleeson, T., Rosolem, R., Pianosi, F., Wada, Y. and Wagener, T.: A large-scale simulation model to assess karstic groundwater recharge over Europe and the Mediterranean, Geosci. Model Dev., 8(6), 1729–1746, doi:10.5194/gmd-8-1729-2015, 2015.

Hartmann, A., Gleeson, T., Wada, Y. and Wagener, T.: Enhanced groundwater recharge rates and altered recharge sensitivity to climate variability through subsurface heterogeneity, Proc. Natl. Acad. Sci., 114(11), 2842–2847, doi:10.1073/pnas.1614941114, 2017.

Inselberg, A.: Parallel coordinates: Visual multidimensional geometry and its applications, Springer, New York., 2009.

Jeannin, P.-Y. and Grasso, D. A.: Permeability and hydrodynamic behavior of karstic environment, in Karst Waters Environmental Impact, edited by G. Gunay and A. I. Johnson, pp. 335–342, A.A. Balkema, Rotterdam, Netherlands., 1997.

Kasprzyk, J. R., Nataraj, S., Reed, P. M. and Lempert, R. J.: Many objective robust decision making for complex environmental systems undergoing change, Environ. Model. Softw., 42, 55–71, doi:10.1016/j.envsoft.2012.12.007, 2013.

Klimchouk, A. B. and Ford, D. C.: Types of karst and evolution of hydrogeologic setting, in Speleogenesis. Evolution of Karst Aquifers, edited by A. B. Klimchouk, D. C. Ford, A. Palmer, and W. Dreybrodt, pp. 45–53, National Speleological Society, Huntsville, Alabama, U.S.A., 2000.

Pianosi, F., Sarrazin, F. and Wagener, T.: A Matlab toolbox for Global Sensitivity Analysis, Environ. Model. Softw., 70, 80–85, doi:10.1016/j.envsoft.2015.04.009, 2015.

Pianosi, F., Iwema, J., Rosolem, R. and Wagener, T.: A multimethod Global Sensitivity Analysis approach to support the calibration and evaluation of Land Surface Models, in Sensitivity Analysis in Earth Observation Modelling, edited by G. Petropoulos and P. Srivastava, pp. 125–144, Elsevier Inc., 2017.

Scanlon, B. R., Keese, K. E., Flint, A. L., Flint, L. E., Gaye, C. B., Edmunds, W. M. and Simmers, I.: Global synthesis of groundwater recharge in semiarid and arid regions, Hydrol. Prpcesses, 20(15), 3335–3370, doi:doi.org/10.1002/hyp.6335, 2006.

Shuttleworth, W. J.: Evapotranspiration, in Handbook of Hydrology, edited by D. R. Maidment, p. 4.1-4.53, McGraw-Hill inc., New York., 1993.

Shuttleworth, W. J.: Terrestrial Hydrometeorology, John Wiley & Sons, Ltd, Chichester, UK., 2012.

Wada, Y., Van Beek, L. P. H. and Bierkens, M. F. P.: Nonsustainable groundwater sustaining irrigation: A global assessment, Water Resour. Res., 48(6), W00L06, doi:10.1029/2011WR010562, 2012.

Williams, P. W.: The role of the subcutaneous zone in karst hydrology, J. Hydrol., 61(1–3), 45–67, doi:10.1016/0022-1694(83)90234-2, 1983.

Williams, P. W.: The role of the epikarst in karst and cave hydrogeology : a review, Int. J. Speleol., 37, 1–10, doi:10.5038/1827-806X.37.1.1, 2008.

Response to Reviewer #2

OVERALL REPLY: We thank the Reviewer for the detailed review and take the opportunity to clarify some important points of our approach, while responding to the criticisms made. We provide detailed responses to specific comments below, but we thought it helpful to summarise the key point here: our manuscript introduces a new evapotranspiration component into a previously developed large-scale hydrological model for karst areas (Hartmann et al., 2015, GMD). We are therefore following an approach to large-scale hydrological modelling that has been widely used for climate change impact studies (e.g. Beyene et al., 2010; Sperna Weiland et al., 2012; Gosling et al., 2017). The Reviewer focuses on the use of land surface models to the same issue, which is an alternative approach currently taken to simulate climate and land cover change impacts on hydrological variables. However, we are not attempting to build a land surface model here, but rather to advance a hydrological model for large-scale applications. Indeed, we ourselves have argued in the past that these two communities should come closer together and learn from each other (Archfield et al., 2015, WRR).
We provide below a point-by-point reply to the comments of the Reviewer with additional references to previous studies that help us support our modelling choices (see list of references at the end of this report). **Line numbers in our replies refer to the marked-up version of our revised manuscript** (included in this file). In the marked-up version of the manuscript we highlighted in red the insertion and deletion we made. We found it difficult to address some of the Reviewers' specific comments, because he/she did not provide any reference to support his/her statements.
In particular, the revised version of the manuscript includes a comparison between hourly and daily simulations of the V2Karst model as discussed in REPLY 1 and 3 below.

COMMENT 1*: The paper proposed by Sarrazin et al. aims at adding a new evaporation formulation to the recharge model VarKarst which specialises on the hydrology of karst systems. The aim of this development is to make the model suitable for exploring the impact of climate and land surface changes on these very sensitive hydrological structures. The main themes of these improvements are to be applicable at the large scale and to be parsimonious.*
I believe this model fails on both accounts for a simple reason, the authors have neglected the fact that evaporation is strongly controlled by the diurnal cycle of radiation and atmospheric processes. One of the main consequences of climate change is to modify the diurnal cycle at the surface and in the atmosphere. Thus the application of V2Karst to climate change is bound to produce unrealistic sensitivities. The model would be more parsimonious and more robust (because based on stronger physical grounds) if it would explicitly represent the diurnal cycle.

REPLY 1: Our model is in line with widely published approaches to climate change assessment using large-scale hydrological models (e.g. Beyene et al., 2010; Sperna Weiland et al., 2012; Gosling et al., 2017), that all neglect the diurnal cycle and apply hydrological models at a daily time step. We provide further details on this issue in the reply to the reviewer's COMMENT 3 (Penman-Monteith equation) and 9 (canopy interception). We agree that there is indeed a strong need for better comparison studies to understand how neglected processes (e.g. of diurnal cycles) affect climate change impact studies, but this is beyond the scope of the study presented here.

Moreover, the reviewer wrote *"One of the main consequences of climate change is to modify the diurnal cycle"*. We agree that global average radiative forcing is projected to increase (IPCC, 2013; Van Vuuren et al., 2011), and average land surface air temperature is documented to have already increased globally (IPCC, 2013, Chapter 2 p187-188 for a global assessment, 2014, Chapter 23 p1275-1276 for an assessment for Europe) and is projected to further increase (IPCC, 2013, Chapter 12 p1062-1064 for a global assessment, 2014, Chapter 23 p1276 for an assessment for Europe). However, to our knowledge, changes in the diurnal cycle appear to be much more uncertain. Changes in temperature are more documented, possibly because the historical temperature record is more accurate than the other climate variables (Allen and Ingram, 2002). We are aware of multiple studies that analysed the past changes in diurnal temperature range (i.e. difference between minimum and maximum daily temperature). A summary of these studies is provided in the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC, 2013). They overall indicate that decreases in diurnal temperature range have been observed, but these decreases were found to be smaller than changes in average temperature (IPCC, 2013, Chapter 2, p188). Moreover, some studies suggest that these apparent changes in diurnal temperature range may be attributed to non-climatic factors (IPCC, 2013, Chapter 2, p188). In this regard, a more recent study points out that the drivers of the past observed changes in diurnal temperature ranges are still not well understood (Davy et al., 2017). Additionally, it has been shown that climate models involved in the Coupled Model Intercomparison Project Phase 5 (CMIP5, scientific basis for the IPCC fifth Assessment Report) are reproducing poorly the past observed changes in diurnal temperature range (Lewis and Karoly, 2013), which suggests that future projections in diurnal temperature range have large uncertainties.

Changes in the revised version of the manuscript

To prepare the revised version of our manuscript, we have performed additional simulations of V2Karst using an hourly time step. We compare the results of the parameter estimation strategy (described in Sect. 4.1) obtained for daily and hourly simulation. We briefly discuss the results in the manuscript p9 L15-19, while detailed results are reported in Sect. S8 of our Supplementary material. We did not find significant differences between daily and hourly simulations for assessing recharge at monthly and annual time scale, which is the focus of our study. This demonstrates that it is reasonable to apply V2Karst at daily time step for our application.

To be able to run the model at hourly time step, we modified the canopy interception routine of V2Karst. We introduced a state variable that represents the interception storage so that the model accounts for the carry-over of interception storage from one time step to the next for hourly time step. The revised interception routine is described p12-13 in the revised manuscript. We have also added the ground heat flux in the description of the Penman-Monteith equation (Eq.(14) p15). For daily simulations, ground heat flux is neglected, while it is accounted for hourly simulations as explained p15 L15-17 and in Sect. S8 of our Supplementary material.

COMMENT 2*:Furthermore this enhancement of VarKarst neglects 30 years in the developments of land surface models. These models do not represent hydrological processes and even less karst systems, and are rightfully criticized for this. But they have specialized on the surface/atmosphere exchanges and in particular the simulation of evaporation, vegetation processes and infiltration. At no moment do the authors refer to developments in one of the three leading land surface models (JULES, ORCHIDEE and CLM) or their application to the 4 FLUXNET stations used here. A simple Google search would have shown to the authors that these open-access codes (Thanks in great part to GMD !) perform much better on these sites and do not require the tuning of so many parameters. Furthermore they are designed to be applicable at the large scale. I would recommend to reject the paper and encourage the authors to download one of the above mentioned land surface models and couple it to VarKarst. This would produce a model for these sensitive hydrological regions which is much more robust and produces more credible result for the impact of climate and land-cover changes. I am sorry to have to make such a harsh recommendation to GMD and in the following I will detail where I believe the basic assumptions of the authors to be wrong and where the usage of developments made for land surface models would help.*

REPLY 2: For clarity, we structured our reply in three parts: (1) we explain why we chose to develop V2karst as an evolution of a parsimonious hydrological model, rather than using more complex land surface models, (2) we attempt to compare the performance of V2Karst and land surface models at the four FLUXNET sites used in our study and (3) we specify the changes that we introduced in the revised version of the manuscript to clarify these points. We do not discuss V2karst representations here, as we will explain them more specifically in REPLY 3 (Penman-Monteith equation), REPLY 6 (soil water balance), and REPLY 9 (interception).

1. Reasons for developing a parsimonious hydrological model

Land Surface Models are undeniably crucial tools because they include state-of-the-art scientific understanding of moisture and energy processes. However, parallel to the development of land surface models, a wealth of studies have drawn attention to the problem of dealing with model complexity in the context of natural systems. In fact, in natural systems, controlled experimentation to ascertain model formulations is not possible and model components and parameters tend to be poorly defined, especially at large-scales, because of a lack of data and knowledge (e.g. Young et al., 1996; Abramowitz et al., 2008; Beven and Cloke, 2012; IPCC, 2013, Chapter 9, pp790-791; Hong et al., 2017; Haughton et al., 2018). For example, Beven and Cloke (2012) highlight the fact that more complex models do not necessarily produce more robust predictions, presumably because of our lack of knowledge of natural processes and because of the uncertainty in estimates/observations of the variables needed to run such complex models. Therefore, regarding future modelling challenges, Beven and Cloke (2012) have argued that understanding which parameterisation may be more appropriate and assessing model uncertainties may be more relevant than further increasing the detail of process representation.

A recent study published in GMD (Haughton et al., 2018) argues that: "*In general, numerical LSMs* [Land Surface Models] *have become increasingly complex over the last 5 decades, expanding from basic bucket schemes to models that include tens or even hundreds of processes involving multiple components of the soil, biosphere, and within-canopy atmosphere. Model components may have been added on to existing models without adequate constraint on component parameters (Abramowitz, 2013) or without adequate system closure (Batty and Torrens, 2001). New component parameters may be calibrated against existing model components, leading to problems of equifinality (Medlyn et al., 2005), non-identifiability (Kavetski and Clark, 2011), and epistemological holism (Lenhard and Winsberg, 2010). These problems can often only be overcome by ensuring that each component is itself well constrained by data and numerically stable. As noted earlier, these conditions rarely exist for any given component.*" In fact, although land surface models have a strong physical basis, they also include many empirical functions that are typically difficult to constrain (Mendoza et al., 2015). More critically, the fact that land surface models include a large number of parameters (many of which hard-coded, as highlighted in Cuntz et al. (2016) or Mendoza et al. (2015)) hampers an exhaustive assessment of uncertainty and sensitivity of model predictions (Young et al., 1996). We further discuss the issue of parameter estimation in land surface models in REPLY 15. Moreover, the study by Haughton et al. (2018, GMD) shows that land surface models can be outperformed by simple empirical models, in line with the results of previous studies conducted within the Land Surface Model Benchmarking Evaluation Project (PLUMBER, Best et al., 2015).

Importantly, land surface models simulate a large range of different fluxes (e.g. sensible heat, latent heat, ground heat flux, radiation, runoff, CO2) and the sensitivity of model parameters has been reported to vary depending on the simulated flux considered (Cuntz et al., 2016; Rosero et al., 2010; Rosolem et al., 2012). For instance, Cuntz et al. (2016) showed that a large number of parameters of the Noah land surface model are non-influential or have a very small influence on total simulated runoff. This means that parts of the land surface models may be simplified when the objective is to simulate hydrological variables. In this sense, Hong et al. (2017) highlighted the fact that model development should account for the model intended uses.

V2Karst aims to simulate seasonal and annual groundwater recharge, because these variables are appropriate to characterise the amount of renewable groundwater, and hence the amount of groundwater available to human consumption and ecosystems (e.g. Scanlon et al., 2006; Döll and Fiedler, 2008; Wada et al., 2012). As generally done in hydrological models (Table A1 of our manuscript), V2Karst focuses on solving the water balance, while it does not solve the energy balance, as land surface models do. In fact, V2Karst is not meant to be used for assessing the energy fluxes (radiation, sensible heat, latent heat and ground heat flux). An additional motivation for us not to solve the energy balance is that its inclusion increases model complexity and computational cost tremendously, which makes it difficult to perform a full uncertainty and sensitivity analysis to assess the adequateness of the different model components. By focusing on the water balance instead and by using parsimonious representations, we enable all components of V2Karst to be subject to uncertainty and sensitivity analysis, as we have explained in Sect. 2.1 p6 L20-30 in our manuscript.

2. Performance of Land Surface Models at the four FLUXNET sites

We are not aware of studies that would allow us to directly infer that JULES, ORCHIDEE and CLM have better performance compared to V2Karst at the four FLUXNET sites, as stated by the reviewer. We found six studies in which the JULES, ORCHIDEE, CLM or Noah land surface models were applied at some of the FLUXNET sites we have used (Anav et al., 2010; Davin et al., 2011; Zhao et al., 2012; Kuppel et al., 2012; Van den Hoof et al., 2013; Chaney et al., 2016). In the following paragraphs, we explain in detail that either the results of these studies cannot be compared to our results, or that the performance of the land surface models appears to be similar to or slightly inferior to those of V2Karst.

The results of four of these studies cannot be compared with our results. The study by Chaney et al. (2016) does not specifically report performance results for any of the four FLUXNET sites. The study by Davin et al. (2011) analysed a full Regional Climate Model including the CLM model and did not present performance results regarding latent heat and soil moisture simulations, which are the two variables we analysed in our manuscript. The study by Kuppel et al. (2012) did not analyse the bias or the correlation coefficient between measured and simulated latent heat/evapotranspiration, while these two metrics were used in our analyses (Sect 4.1 of our manuscript). In fact, both metrics are important to characterise the hydrological performance of the recharge model, since the bias assesses the performance in reproducing the overall water balance, while the correlation coefficient assesses the consistency between the temporal pattern of simulated and observed latent heat/evapotranspiration. Finally, the study by Anav et al. (2010) did not report quantitative performance metrics for the individual sites.

In the two "comparable" studies (Zhao et al., 2012; Van den Hoof et al., 2013), land surface models show similar or slightly inferior performance compared to V2Karst, although it is difficult to make a fair comparison because the time periods analysed were not the same as in our manuscript. Zhao et al. (2012) assessed the squared correlation coefficient R^2 between monthly observed and simulated latent heat at the French 2 site (Puéchabon) using the ORCHIDEE model and local forcing data. They found a value of R^2 of 0.59, i.e. a correlation coefficient of $\sqrt{0.59} = 0.77$ (Zhao et al., 2012, fourth row in Figure 6). This is comparable to the performance of the V2Karst model, since we identified simulations for which the correlation coefficient between monthly simulated and observed ET was higher than 0.77 (Figure 4 in our manuscript). We wish to mention that there appears to be a mistake in the labels in Figure 6 in Zhao et al. (2012) and therefore we are not completely sure we have read the figure correctly.

Finally, in Van den Hoof et al. (2013), the JULES model was tested at the German site (Hainich) using local forcing data. The bias between observed and simulated evapotranspiration for the best performing version of the model is around 29% (this percentage is inferred from Figure 2.b in Van den Hoof et al. (2013), which reports that the mean annual observed ET is 280 mm.y^{-1} and the simulated mean annual ET is 360 mm.y^{-1}). The bias is therefore above the limit of acceptance of 20% that we have used in our manuscript. Van den Hoof (2013) did not provide the monthly correlation values for individual sites, but we can infer that, for the German site, the monthly correlation coefficient is between 0.7 and 1 (Van den Hoof et al., 2013, Figure 5.b), which is comparable to our study.

3. Changes in the revised version of the manuscript

Based on the Reviewer's comment, we realise that in our manuscript we have not stated clearly enough our philosophy in developing V2Karst in line with other large-scale hydrological models. We clarified this point in the introduction p5 L13-19 (in the paragraph that states the objectives of the study). We also added further justifications and references to support our choice of developing a parsimonious model in Sect. 2.1 p6 L20-30.

We agree that for completeness our manuscript should also refer to the literature on land surface modelling. In the revised version of our manuscript we include a discussion of the ET components of land surface models in Sect. 2.1 p7 L15-21. We also explain in this paragraph why we chose to develop a hydrological model.

The paragraph in the conclusions section p34 L16-25 that was already present in the previous version of our manuscript summarises our modelling strategy.

However, we would avoid including a comparison between the performance of the V2Karst model and other land surface models at FLUXNET sites because, as shown in our point (2) above, it is difficult to make a fair comparison between different studies that used different metrics and different time periods to evaluate the models.

COMMENT 3: *Rational to explicitly represent land cover properties :*
It is laudable for the authors to use the Penman-Monteith formulation for potential evaporation. But should they have paid attention to its derivation, they would have noted that it only provides potential evaporation over a infinitesimal time intervals as it assumes that atmospheric variables and surface states do not evolve through other processes. A constant Rn(t) or ra,can(t) over the course of the day is a very unsatisfactory assumption, especially under a changing climate. Because of very contrasted impact of changing atmospheric composition on long-wave and short-wave radiation, we can encounter the same Rn but with very different radiation balance, turbulent fluxes and surface temperatures. The authors will find in the literature a number of paper which examine the impact of climate change on the different potential evaporation formulation. They all recommend to use sub-diurnal solutions because of the modified diurnal dynamics.

REPLY 3: We are aware of the fact that the Penman-Monteith equation has been determined over an infinitesimal time step (e.g. Shuttleworth, 2012) and that, therefore, using sub-daily rather than daily time step for the calculation of the equation provides conditions that are closer to its theoretical derivation. Nevertheless, Penman-Monteith equation has been shown to be applicable at daily time step and is widely used at daily time step.

In particular, Allen et al. (2006) commented on applications of the Penman-Monteith reference crop framework, that was designed by the Food and Agriculture Organization of the United Nations (FAO) and the American Society of Civil Engineers (ASCE). The framework has standardised the use of the Penman-Monteith equation for both hourly and daily time step (Allen et al., 1998; Walter et al., 2005) with a focus on agricultural crop. Allen et al. (2006) stated the following: "*The favorable performance* *of the PM equation in many studies, when applied with 24-h (and even monthly) time steps, is somewhat surprising, since the formulation of the combination equation (combined energy balance and aerodynamic components) theoretically requires weather inputs on a nearly instantaneous basis. The general consistency and accuracy of the PM method for 24-h time steps speaks to the combination equation's robustness in estimating evaporative behavior given a particular set of meteorological conditions.*" More recent papers still support both hourly and daily applications of the Penman-Monteith reference crop framework (e.g. Pereira et al., 2015). The daily framework has been included for instance in an R package for simulating evapotranspiration at daily time step (Guo et al., 2016).

The Penman-Monteith equation is widely used at daily time step. Among its applications, we can cite the Moderate Resolution Imaging Spectroradiometer (MODIS) Evapotranspiration product (Mu et al., 2011; Running et al., 2017). Furthermore, the Penman-Monteith equation is implemented in the large-scale hydrological models PCR-GLOBWB, MacPDM and VIC (these models are reviewed in Table A1-A3 of our manuscript). These models have been applied at daily time step in climate change impact studies. For instance, Gosling and Arnell (2016) used MacPDM, Sperna Weiland et al. (2012) used PCR-GLOBWB and Beyene et al. (2010) used VIC at daily time step. These three hydrological models are also included in the Inter-Sectoral Impact Model Intercomparison Project (ISI–MIP, Warszawski et al., 2014) that aims to assess the impact of climate change. In that context, daily simulations of PCR-GLOBWB, MacPDM and VIC are analysed for instance by Gosling et al. (2017).

Changes in the revised version of the manuscript

Given all the above, we think that applying the Penman-Monteith equation at daily time step is one possible choice that is consistent with a wide range of published studies. Additionally, our comparison between hourly and daily time step performed in the revised version of the manuscript (as explained in REPLY 1) demonstrates that it is reasonable to apply V2Karst at daily step for our application. This comparison is briefly discussed in the manuscript p9 L15-19, while detailed results are reported in Sect. S8 of our Supplementary material.

Additionally, our manuscript refers to other hydrological models that applied the Penman-Monteith equation at a daily time (Table A2). In the revised version of the manuscript, when discussing the Penman-Monteith equation, we added a reference to Allen et al. (2006) and Pereira et al. (2015) to further justify the use of the Penman-Monteith equation at both sub-daily and daily time step p15 L4.

COMMENT 4*: The parsimony of our representation of nature if not for us to choose. We have to prove that certain simplifications in the representation of surface processes are valid for the application we envisage. The authors aim to develop a model valid at the large scale, for climate and land surface change. Is it then reasonable to assume that over the course of a day ra,can does not change ? I think the development of land surface models has shown that one cannot neglect the diurnal dynamic of the opening of the stomata, the soil moisture stress or the dependence of stomatal resistance to atmospheric CO2 concentration. If the authors believe that they have found a way to represent with a single daily value these complex processes and their interaction with the environment they should let the world know as it would allow land surface models to be simplified.*

REPLY 4: We explain in REPLY 1 and 2 why we have chosen to include a parsimonious process representation in V2Karst, which is in line with well-established large-scale hydrological modelling studies on climate change impacts. An in-depth comparison with land surface models would be interesting, but it is beyond the scope of our work here.

The evaluation of the Penman-Monteith equation, which requires the specification of the aerodynamic resistance ra_can, is well documented at a daily times step (see REPLY 3).

Our representation does include soil moisture stress using a linear function of soil moisture which multiplies the potential evapotranspiration rate (Eq.(11-12) p14 of our revised manuscript). This formulation is similar to the previous version of the model (VarKarst, Hartmann et al., 2015) and to other hydrological models (Table A2 of our manuscript).

V2Karst neglects the dependence of the stomatal resistance on atmospheric factors such as temperature, radiation, humidity and CO2 concentration, as implemented for instance in Ball-Berry (Sellers et al., 1996) and Jarvis-Stewart (Jarvis, 1976; Stewart, 1988) schemes. Stomatal resistance is assessed as a function of a minimum stomatal resistance (which is a parameter of the model) and the leaf area index (Eq. (15) p16). This formulation is similar to the hydrological model MacPDM (Table A2) that has been applied in climate impact studies (e.g. Gosling and Arnell, 2016) as already reported in REPLY 3. Compared to MacPDM, V2Karst also includes the dependence of the stomatal resistance on the leaf area index, although we recognise that this approach is still a simplification of stomatal processes. The reason for our choice comes from the fact that the value of the stomatal resistance is poorly characterised for large-scale applications, because few ground measurements of stomatal resistance are available (some are reported in Breuer et al. (2003) and Körner et al. (1995)). Moreover, the temporal variability of stomatal resistance makes its measurements particularly difficult to interpret (Breuer et al., 2003) and therefore to use in modelling applications. We have discussed the lack of ground measurements of stomatal resistance (and more generally of vegetation properties) in Sect. 2.1 p6 L2-7 and in Sect S1 of our Supplementary material. The Ball-Berry and Jarvis-Stewart parameterisations of stomatal resistance significantly increase model complexity and in particular introduce several empirical parameters, whose values are uncertain because of the lack of stomatal resistance measurements. The constant value of the stomatal resistance implemented in V2Karst allows us to lump all the uncertainties around stomatal resistance estimation into a single parameter, that we can include in our uncertainty and sensitivity analysis. In this way, we can easily analyse the impact of the value of the stomatal resistance on simulated recharge. We obviously do not claim to "*have found a way to represent with a single daily value these complex processes and their interaction with the environment*", but simply to have included in our model a simple approach to control the overall impact of this variable on recharge predictions.

The temporal behaviour of the stomatal resistance could be investigated in V2Karst in future studies. However, we believe that, as a first version of the model, such in-depth investigation is beyond the scope of this study.

COMMENT 5*: In their rational for their modelling strategy they only mention one land surface model : ISBA in its 1998 version. This is not up to date. Even ISBA has evolved since then and does not use any more a Jarvis type parametrisation. It now also uses a Ball-Berry type formulation which balances carbon uptake and transpiration. Please note that ISBA operates at sub-diurnal time steps.*

REPLY 5: The Reviewer is right, the description of the ISBA model is not up-to-date. Moreover, since the ISBA and LaD are land surface models, while we follow an approach to large-scale hydrological modelling, we deleted the reference to ISBA and LaD in Table A1-A3 in the revised manuscript. The revised version of the manuscript includes a discussion of the ET components of land surface models in Sect. 2.1 p7 L15-21 as already stated in REPLY 2. We explain in this paragraph why we chose to develop a hydrological model.

COMMENT 6*: Soil water balance :*
The explanations of the evolution of moisture in the unsaturated zone is not very clear to me. It looks to me like a superposition of buckets with the addition of lateral flows. It has been the experience in the land surface model community that this simple representation of soil moisture limits the ability to simulate the impact of

stresses on transpiration. This is particularly critical in semi-arid those encountered at 3 of the selected FLUXNET stations. What is the reasoning of the authors behind this simplification in the treatment of the unsaturated zone, apart from "parsimony" ?

REPLY 6: The evolution of moisture in the saturated zone in V2Karst is the same as in VarKarst (Hartmann et al., 2015, GMD). This representation is supported by physical reasoning and by previous testing of the VarKarst model.

As explained in Sect. 2.3.2 p10-11, percolation from a given soil layer (Soi_i) to the underlying soil layer (Soi_k) is equal to the saturation excess in Soi_i. Percolation from the deeper (third) soil layer to the epikarst layer is also equal to the third soil layer saturation excess. As explained p3 L10 and p11 L6-12, the reasoning behind this simplification is that soils in karst areas typically have a high clay content (Blume et al., 2010), and therefore tend to have low unsaturated permeability (Clapp and Hornberger, 1978). Instead, saturated permeability is typically high because clay soil generally have cracks that can act as preferential flow pathways under wet conditions (Lu et al., 2016).

The representation of the process of lateral flow is a feature specific to the VarKarst model, which, together with the explicit representation of sub-surface sub-grid heterogeneity (using model vertical compartments), makes the model more appropriate for karst areas than other large models. The adequacy of the soil water balance representation in VarKarst has been tested in previous studies for previous versions of the model (Hartmann et al., 2012, 2015, 2017). In particular, Hartmann et al. (2015, 2017) showed that there is no systematic bias in the model predictions, and that recharge predictions produced by VarKarst are significantly higher than those produced by models that do not include karst processes. Further discussion of the representation of diffuse and concentrated flow in V2Karst and of the testing of the previous versions of the model are given in our response to the first review of this manuscript (REPLY 2 in 'Response to Reviewer #1').

Changes in the revised version of the manuscript

In the revised version of our manuscript, we added a briefly explanation on how diffuse and concentrated recharge is represented in VarKarst and V2karst in Sect. 2.2. p8 L12-17 which was missing.

COMMENT 7*: Evapotranspiration :*

Only one vegetation type seems to be allowed per grid-box, is this correct ? Because of the strong heterogeneity of the distribution of vegetation, it has been the experience of the community that a larger number of plant functional types is needed per grid box. The strict minimum has been found to be a low and a high vegetation. This simplification will be critical for the application to larger domains and in particular in semi-arid regions where the competition of the various vegetation types for water is critical.

REPLY 7: We agree that, for large-scale applications, the model should be able to represent different vegetation types within a given simulation grid cell. The model can account for sub-grid heterogeneity in vegetation type using a 'tile' approach.

A 'tile' approach consists of subdividing each model grid cell in a number of independent units (tiles), each of which has a specific land cover (e.g. short or tall vegetation). The model can then be evaluated separately over each tile. The overall simulated fluxes for a given grid cell are computed as the area weighted average of the fluxes calculated over the tiles. The same approach is also used in other large-scale hydrological models, for instance in the Mac-PDM model (Gosling and Arnell, 2011) and in the VIC model (Bohn and Vivoni, 2016; Liang et al., 1994, http://vic.readthedocs.io/en/master/Overview/ModelOverview/).

Changes in the revised version of the manuscript
We added a paragraph in Sect. 6.3 p33 L10-14 (this section discusses large-scale application of V2Karst) to clarify the fact that sub-grid heterogeneity in vegetation types can be treated following a tile approach.

COMMENT 8*: Please explain here as well why the literature on vegetation modelling is not relevant for this model.*

REPLY 8: We are not sure which specific literature the reviewer is referring to. In Sect. 2.1 and Tables A1-A3 of our manuscript we have provided a detailed review of ET and vegetation modelling in other large-scale hydrological models. As mentioned in REPLY 2 and 5, we also added a brief discussion of the ET components of land surface models in Sect. 2.1 of our revised manuscript.
We are aware of the fact that our manuscript is rather long and we have already been advised by another reviewer to shorten it, hence we would avoid adding further details on vegetation modelling besides Sect. 2.1.

COMMENT 9*: Canopy interception :*
This is another topic where the community has acquired a rich experience which could benefit the authors. The representation of canopy interception at different temporal and spatial scales has been fiercely debated in the early 90s. Thus a number of parametrisations were developed to take into account the spatial and temporal variability of interception. This would be relevant here. Do the authors believe that a rainfall event in the evening or in the morning produces the same interception loss ? Does a rainfall event of 10mm/h and 100mm/h produce the same interception ? So does the assumption of treating these processes averaged over the day have any implication on the sensitivity of V2Karst to climate change ? We know that rainfall intensity and possibly also the time of day at which precipitation will occur will change in a warmer climate.

REPLY 9*:* V2Karst implements a parsimonious daily interception model as done in other large-scale hydrological models.

In V2Karst, for each day, evaporation from the interception store is assessed over the vegetated fraction as the maximum value between daily precipitation, potential evapotranspiration and interception storage capacity. Our model is similar to the large-scale hydrological models WaterGap, PCR-GLOBWB and VIC (Table A3 in our manuscript). These models have been simulated at daily step to assess the hydrological impact of future changes in climate, for instance in Gosling et al. (2017) within the Inter-Sectoral Impact Model Intercomparison Project (ISI–MIP, Warszawski et al., 2014), and in Döll et al. (2018) for WaterGap, in Sperna Weiland et al. (2012) for PCR-GLOBWB and Beyene et al. (2010) for VIC. Past studies have shown that interception can be reasonably represented at a daily time scale (Savenije, 1997; De Groen and Savenije, 2006; Gerrits et al., 2009).

We are aware of the fact that more sophisticated interception schemes have been developed, as reviewed for instance in (Muzylo et al., 2009). These more complex schemes include additional parameters compared to our formulation that includes only one parameter, namely the interception storage capacity. We chose to implement a simple interception scheme for large-scale applications because (1) the physical processes and atmospheric conditions driving evaporation from canopy interception are still poorly understood (Van Dijk et al., 2015), (2) ground measurements of evaporation from interception that would be necessary to constrain model parameters are not available (Fatichi and Pappas, 2017; Miralles et al., 2016), and (3) ground measurements of canopy interception are also affected by large uncertainties (Van Dijk et al., 2015).

As already mentioned in REPLY 4, we have discussed the issues arising from lack of measurements in large-scale model applications in our model development rationale (Sect. 2.1 p6 L2-7 and in Sect S1 of our supplementary material). Similar to our representation of stomatal resistance (REPLY 4), since the representation of interception processes can only be poorly constrained for large-scale applications, we chose a simple interception scheme. In this way, we can easily include the uncertainty in interception representation in our analyses and gain a systematic understanding of the impact of this uncertainty on seasonal and annual recharge predictions (our variables of interest).

Changes in the revised version of the manuscript

Additionally, as explained in REPLY 1, the revised version of the manuscript includes a state variable that represents the interception storage so that V2Karst accounts for the carry-over of interception storage from one time step to the next for hourly time step. The revised interception routine is described p12-13 in the revised manuscript.

COMMENT 10*: May I point out at this stage that precipitation intensification has been observed at the sub-diurnal range. Daily mean rainfall has not yet been too much affected by climate change. On the other hand, hourly precipitation rates have been increasing faster than expected from the Clausius-Capleyron relation. Thus, the virtual experiments experiments proposed in section 4.3 are not relevant for climate change. The authors are referred to the wealth of literature published on this topic in the last few years.*

REPLY 10*:* We have actually found evidence suggesting that daily mean rainfall has been affected by climate change, and therefore analysing the impact of changes in daily precipitation (overall mean and temporal distribution), as done in our virtual experiment, is relevant.
The last IPCC assessment reported that many regions show a significant trend (positive or negative) in mean daily precipitation for the past period (IPCC, 2013 and particularly Fig. 2.33, reported as Figure R1 below). Additionally, a more recent study by Ye et al. (2016) found that an increase in mean daily precipitation intensity occurred in all seasons in the northern Eurasia region over the period 1966-2010. Regarding future climate, model projections indicate changes in seasonal precipitation and changes in the 95 percentile of daily precipitation, for instance in Europe (IPCC, 2014, Chapter 23, Section 23.2.2.2 and Fig. 23-2).

Changes in the revised version of the manuscript

In the revised version of the manuscript, we clarify the fact that daily precipitation is likely to change in Sect. 4.3 p22 L4.

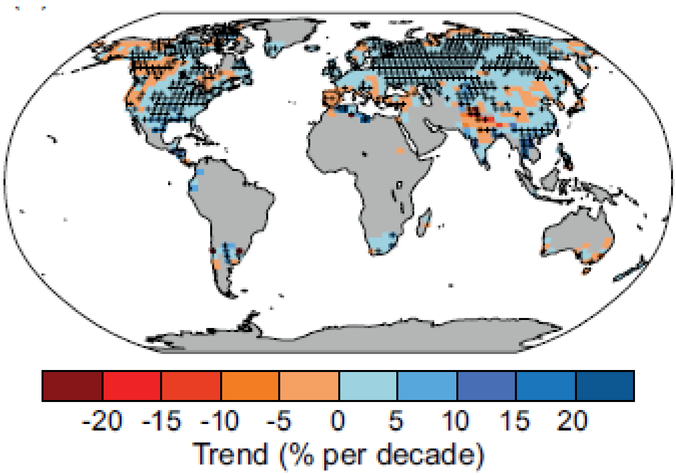


Figure R1. Trend in daily precipitation intensity over the period 1951-2010. Trends were calculated only for grid boxes that had at least 40 years of data during this period and where data ended no earlier than 2003. Grey areas indicate incomplete or missing data. Black plus signs (+) indicate grid boxes where trends are significant (i.e., a trend of zero lies outside the 90% confidence interval).
Source of the Figure: (IPCC, 2013, Figure 2.33)

COMMENT 11: *Transpiration from vegetated soil :*
Transpiration does not occur from the soils (as written in the paper) but from the stomata in the vegetation. This is not a negligible detail. Firstly the stomata only open when daylight is present and thus photosynthesis can occur. During the early afternoon, once the water in contact with the roots and within the plant has been evaporated, transpiration declines. This is caused by the slower diffusion of water within the soil which limits the supply. This is known to be a critical process for transpiration and which will be affected by higher CO2 concentration which will lead plants to reduce the opening of their stomata. I guess these processes are neglected in the proposed model, why ? It would be a very interesting topic to see how this early afternoon depression of transpiration is affected by climate change for plants on karstic soils. It is bound to be different than on loamy soils for instance.

REPLY 11: We agree with the reviewer that "*transpiration does not occur from the soil*". In the revised version of the manuscript, we will correct this mistake. More specifically, p13 L19 we replaced 'Transpiration from vegetated soil' by 'Transpiration over the vegetated fraction'.

Regarding the reviewer's comment "*I guess these processes are neglected in the proposed model, why ?"*, as explained in detail in REPLY 4, we neglect the dynamics of stomatal resistance because lack of observations does not allow constraining complex stomatal resistance schemes at a large-scale. We instead use a constant value of the stomatal resistance (scaled by the leaf area index). This enables us to easily analyse the effect of uncertain stomatal resistance on simulated recharge by including the stomatal resistance parameter in our sensitivity analysis.

"*How this early afternoon depression of transpiration is affected by climate change for plants on karstic soils*" is a very interesting research question but is beyond the scope of our research. Again, the aim of our modelling activity here is to (reasonably, given all data and knowledge limitations) assess the effects of climate and land cover changes on groundwater recharge at the spatial and temporal scale of interest (large/regional; seasonal/annual). Our objective is not to assess the effect of climate variations on every specific process involved in plant transpiration. As stated in REPLY 2, we have clarified the objectives of our study p5 L13-19.

COMMENT 12: *Sorry, the assumption "... evaporation from interception is constant throughout the day ..." is not valid and will change with climate and land surface type.*

REPLY 12: We agree with the reviewer that the justification of Eq. (13) p14 (fraction of the day with wet canopy) has been expressed incorrectly, ultimately misleading the interpretation of our proposed approach.

We do not actually make the strong assumption that "evaporation from interception is constant throughout the day". We instead assess the fraction of the day with wet canopy as the fraction of available energy that was used to evaporate water from the interception store. This conceptual representation was proposed in Kergoat et al. (1998) and it was adopted in the dynamic vegetation model LPJ (Gerten et al., 2004; Murray, 2014) presented in Table A1-A3. In V2Karst, we revised this formulation, which in the referenced studies uses the fraction of daytime with wet canopy (hence assuming that ET fluxes occur during daytime only), by using the fraction of the entire day, given that night time ET fluxes have been shown to be also important (e.g. Pearce et al., 1980; Sugita and Brutsaert, 1991; Kelliher et al., 1992).

Changes in the revised version of the manuscript

In the revised version of the manuscript, we clarified and replaced the statement the point by replacing the statement p14 L12-15 ("The fraction of the day with wet canopy $t_{wet}(\mathrm{t})\ [-]$ is estimated by assuming that evaporation from interception is constant throughout the day ...") by "The fraction of the day with wet canopy $t_{wet}(\mathrm{t})\ [-]$ is estimated as the fraction of available energy that was used to evaporate water from the interception store …".

COMMENT 13: *Parameter estimation :*

The proposed parameter estimation is difficult to interpret in view of the strong hypothesis made in the basic equations of the model. The 15 parameters of this model are so conceptual, i.e. far away from first physical principles, that indeed they can all be tuned. But given the large number of "tunable parameters" can it not be expected that the model can be made to match any dataset ? To me hveg, LAI(min,max) or z0 are not "tunable parameters" as they can either be measured or derived from turbulence theory.

REPLY 13: We agree with the Reviewer that (1) the ability to identify parameter sets that produce predictions close to observations is a necessary but not sufficient condition for the plausibility of the V2Karst model and (2) some parameters of V2Karst can be derived from a priori information. We think that we have accounted for these two points in our methodology, as we explain below.

Regarding (1), we agree that a model can produce reasonable predictions by 'incorrectly' activating processes to compensate for structural deficiencies. For this reason, in our evaluation of V2Karst, in addition to the comparison between observations and predictions (Sect. 4.1), we also performed two sensitivity analyses: one using measured forcing data (Sect. 4.2) and the other one using synthetic forcing data (Sect. 4.3). The aim of these analyses was precisely to assess whether the sensitivities of V2Karst outputs across sites are consistent with our understanding of the physical characteristics of those sites, and hence whether the model reproduces the observations for the right reasons by activating the appropriate controlling factors at the right place. In our manuscript, we discuss the fact that the global sensitivity analysis applied to V2Karst showed a set of sensitivities that are interpretable in light of the different climatic conditions at the four FLUXNET sites (see Sect. 5.2 p26 L27-p27 L2 and Sect. 6.1 p29 L17-27). Such use of sensitivity analysis as a tool for model diagnostic evaluation (verification of model structures) has been successfully applied already in several modelling studies (e.g. Rosero et al., 2010; Reusser and Zehe, 2011; Rosolem et al., 2012; Hartmann et al., 2013).

Regarding (2), we have actually used measurements at the FLUXNET sites to constrain the value of parameters hveg, LAI(min,max), Vr and Vsoi. As explained in Sect. 4.1, our parameter estimation strategy is based on the sequential application of 'soft rules' to identify plausible parameter values within an initial sample spanning over a wide range. Rules 1-4 select 'plausible' parameter sets based on their ability to produce model outputs consistent with observations, while the last rule (rule 5) selects the parameter sets that are consistent with a priori information about the sites. Constraining parameter values based on a priori information in the last step, instead of doing it in the first step as typical of other estimation strategies, provides us with a more stringent test of the adequacy of the model structure. In fact, if the parameter sets identified as plausible by Rules 1-4 (consistency with observations) are also plausible according to Rule 5 (a priori knowledge about the site), we can indeed conclude that the model 'gives the right response for the right reason'. If instead some parameter sets identified as 'plausible' by the comparison with observations (Rules 1-4) are then ruled out when including a priori information (Rule 5), this indicates some deficiencies in the model structure that are being 'compensated for' by tunable parameters and points us to the model components that need improvement. In our experiment, presumably because too few data are available to constrain the simulations, Rules 1-4 result in little constraining of the parameter ranges (Figure 4, Sect. 5.1). However, importantly, we can identify parameter sets that satisfy all rules simultaneously and that are therefore consistent with a priori information at the site, while also producing predictions consistent with the available observations.
A similar strategy was used in previous studies, and more specifically in Hartmann (2015) to estimate the parameters of the VarKarst model and in Rosero et al. (2010) to estimate soil and vegetation parameters of the Noah land surface model.

In future large-scale applications of V2Karst, we envisage that parameters hveg, LAI(min,max), Vr and Vsoi can be a priori constrained based on grid-cell specific information about vegetation and karst landscape type (Vsoi was constrained based on karst landscape type in Hartmann et al. (2015)).

Changes in the revised version of the manuscript
In the revised version of the manuscript, we clarify the meaning and implications of our sensitivity analyses and the strategy we envisage to estimate V2Karst parameters in future large-scale applications. We revised the introduction to the method section (Sect. 4) to make it clearer and more informative (p19 L10-19). We rewrote Sect. 6 more concisely to better highlight the meaning and implications of our sensitivity analyses (Sect. 6.1 and 6.2 p29-32) and the strategy we envisage to estimate V2Karst parameters in future large-scale applications (Sect. 6.3 p32-33). In particular, in Sect. 6.3, we explain that some model parameters will be estimated using a priori information for the different simulation grid-cells (p32 L23-25). Additionally, we rewrote the last paragraph of the conclusions section (p35 L5-11) to stress the importance of using global sensitivity analysis to check that the model is able to reproduce the data for the 'right reasons'.

COMMENT 14: *Furthermore I find that the range of values explored for these parameters (Table 3 does not provide the limits for all 15 parameters) is much wider than realistic values I have observed.*

REPLY 14: Table 1 provides the wide ranges for all 15 parameters that were used to derive the initial parameter sample for our parameter estimation strategy. Table 3 provides the a priori ranges that were used for applying our soft rule 5. These ranges are defined only for those parameters for which site-specific a priori information was available.

All these ranges were determined in accordance with values found in the literature. References and detailed explanations are given in Sect. S3 of our Supplementary material (Table S11 and S12 p14-16).

COMMENT 15: *Land surface models also use the FLUXNET observations to "tune parameters". But fewer parameters are adjusted and only those where the definition itself includes processes which are not modelled, i.e. are conceptual. Furthermore these parameters are specific to the plant functional type present at the FLUXNET station and then then transferred to the larger scale. This is the value of using vegetation classes in land surface models. A simple internet search for FLUXNET and the name of one of the leading land surface models, returns a large number of papers. Some where the models are simply validated and others where the observations are used to refine some vegetation parameters. The authors should have done that search during the development of their model.*

REPLY 15: We are aware of the fact that land surface models indeed typically tune a very limited number of model parameters. However, the results of recent studies suggest that excluding many parameters from calibration and uncertainty/sensitivity analysis raises a number of issues and that parameter estimation strategies of land surface models should be enhanced, as explained in the next paragraphs.

Among the parameters of land surface models that are not 'tuned', some are considered as 'physical' parameters and are commonly read from look-up tables (more specifically soil and vegetation parameters are assigned based on soil texture and vegetation type respectively) or are set to site-specific values when measurements are available. Other parameters, which have an empirical basis, are set to constant values or are even hard-coded, i.e. are embedded in the model code as fixed values (e.g. coefficients to describe a particular shape for a curve in the model, such as the exponent used in the snow depletion curves for the melting season in the Noah land surface model as reported in Mendoza et al., (2015). Hogue et al. (2006) highlighted the fact that many parameters in land surface models cannot be directly obtained from field measurements but through curve-fitting techniques, such as parameters controlling the soil hydraulic properties (derived for instance in Clapp and Hornberger, 1978).

Previous studies suggest that reading parameters from look-up tables or setting them to fixed values are not satisfying strategies. Some studies have put into questions the physical meaning of vegetation and soil parameters in land surface models, and therefore the fact that they can be assigned from look-up tables or from site measurements. Rosero et al. (2010) analysed the sensitivity and optimal values of different vegetation and soil parameters for different version of the Noah land surface model along a precipitation gradient in the southern USA. Their results suggest that climate strongly controls the optimal value of soil and vegetation parameters. This means that vegetation and soil parameters do not strictly represent vegetation and soil properties but also account for other properties. Hogue et al. (2006) analysed the differences in optimal values of different vegetation and soil parameters across five different land surface models (BUCKET, CHASM, BATS1e, BATS2, and Noah) and for five different vegetated sites in the USA. They found significant differences in optimal parameter values for the same parameters and the same site across models. This suggests that the vegetation and soil parameters have a different meaning across models. Therefore, they conclude that land surface models should be interpreted as simplified conceptual representations of natural systems, in which parameters are also conceptual representations of physical properties. The incommensurability between model parameters and physical properties can be explained by the scale mismatch between physical variables that can be actually measured and model parameters that represent cell average properties. An additional possible explanation for this incommensurability is that models implement governing equations that are well established to describe the behaviour of the system over small scales (e.g. Darcy's law), but that may not be valid for applications over larger domains (Kirchner, 2006).

Moreover, it has been shown that some empirical parameters that are typically fixed to constant values, and in particular some hard-coded parameters, have a significant impact on model predictions and therefore these parameters would need to be included in the parameter estimation strategy, i.e. to be 'tuned' (Cuntz et al., 2016; Mendoza et al., 2015).

The fact that land surface models include a large number of parameters that would potentially need to be 'tuned', and the limited availability of data to constrain model predictions at large-scales (as discussed p6 L2-7 of our revised manuscript, in Sect. S1 of our supplementary material, and in REPLY 4 and 9 above) make the calibration of land surface models particularly challenging in large-scale applications. How to enhance parameter estimation of land surface models is indeed still an open question (e.g. Chaney et al., 2016).

Changes in the revised version of the manuscript
In the revised version of our manuscript, we added a discussion on the challenge of estimating the parameters of large-scale models, and in particular land surface models (Sect. 2.1 p6 L2-7 and p7 L17-21).

COMMENT 16*: Conclusion :*
I am very sorry to have to write this review about the development of V2Karst. I know what a huge effort it is to develop a complex numerical model. As the authors are working in Britain, I would recommend that they look into the JULES land surface model. It is freely available and could be coupled to VarKarst to produce a very innovative tool which could indeed allow to explore the consequences of climate and land surface changes on water resources of karst aquifers. This need, to initiate a convergence between hydrological and land surface modelling, has been recognized by NERC and lead to the initiation of the HydoJULES program. The authors should contact the leaders of this program to obtain help.

REPLY 16: The V2karst V1.0 model is an extension of a large-scale hydrological model previously published in GMD (VarKarst, Hartmann et al., 2015) that aims to simulate groundwater recharge in karst areas under changing environmental boundary conditions. Hydrological models are widely applied to study groundwater recharge at large-scales. As we have stressed in REPLY 2 and 15, land surface models simulate many more fluxes than hydrological models (e.g. sensible heat, latent heat, ground heat flux, radiation, CO2), which implies that they include many more parameters that are difficult to constrain in large-scale applications. Since our objective is to predict seasonal and annual recharge (as stated in our manuscript p5 L13-19), which is the key variable of interest for water resources management in karst areas, and not to assess the other fluxes simulated by land surface models, we chose to build on and expand a hydrological model instead of using a land surface model. While we acknowledge that the approach proposed by the Reviewer may be an alternative route to achieve our goal, we hope we have provided a convincing rationale for our modelling choices.
We took the opportunity to revise our manuscript to clarify these points and to assess the effect of the simulation time step on the simulation results (p9 L15-19 and in Sect. S8 of our Supplementary material). We demonstrated that the effect of the simulation time step is not significant for our application and that it is reasonable to run V2Karst using a daily time step.

References to support our reply to Reviewer #2

Abramowitz, G.: Calibration, compensating errors and data-based realism in LSMs, Presentation, 2013.

Abramowitz, G., Leuning, R., Clark, M. and Pitman, A.: Evaluating the performance of land surface Models, J. Clim., 21(21), 5468–5481, doi:10.1175/2008JCLI2378.1, 2008.

Allen, M. R. and Ingram, W. J.: Constraints on future changes in climate and the hydrologic cycle, Nature, 419(6903), 224–232, doi:10.1038/nature01092, 2002.

Allen, R. G., Pereira, L. S., Raes, D. and Smith, M.: Crop evapotranspiration: Guidelines for computing crop requirements, FAO Irrigation and Drainage Paper 56, Food and Agriculture Organization (FAO), Rome, Italy., 1998.

Allen, R. G., Pruitt, W. O., Wright, J. L., Howell, T. A., Ventura, F., Snyder, R., Itenfisu, D., Steduto, P., Berengena, J. and Yrisarry, J. B.: A recommendation on standardized surface resistance for hourly calculation of reference ETo by the FAO56 Penman-Monteith method, Agric. Water Manag., 81(1-2), 1–22, doi:10.1016/j.agwat.2005.03.007, 2006.

Anav, A., D'Andrea, F., Viovy, N. and Vuichard, N.: A validation of heat and carbon fluxes from high-resolution land surface and regional models, J. Geophys. Res. Biogeosciences, 115(G4), 13,281–13,235, doi:10.1029/2009JG001178, 2010.

Archfield, S. A., Clark, M., Arheimer, B., Hay, L. E., McMillan, H., Kiang, J. E., Seibert, J., Hakala, K., Bock, A., Wagener, T., Farmer, W. H., Andréassian, V., Attinger, S., Viglione, A., Knight, R., Markstrom, S. and Over, T.: Accelerating advances in continental domain hydrologic modeling, Water Resour. Res., 51(12), 10078–10091, doi:10.1002/2015WR017498, 2015.

Batty, M. and Torrens, P. M.: Modeling complexity: the limits to prediction, Cybergeo Eur. J. Geogr., doi:10.4000/cybergeo.1035, 2001.

Best, M. J., Abramowitz, G., Johnson, H. R., Pitman, A. J., Balsamo, G., Boone, A., Cuntz, M., Decharme, B., Dirmeyer, P. A., Dong, J., Ek, M., Guo, Z., Haverd, V., van den Hurk, B. J. J., Nearing, G. S., Pak, B., Peters-Lidard, C., Santanello, J. A., Stevens, L. and Vuichard, N.: The Plumbing of Land Surface Models: Benchmarking Model Performance, J. Hydrometeorol., 16(3), 1425–1442, doi:10.1175/JHM-D-14-0158.1, 2015.

Beven, K. J. and Cloke, H. L.: Comment on "hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water" by Eric F. Wood et al., Water Resour. Res., 48(1), W01801, doi:10.1029/2010WR010090, 2012.

Beyene, T., Lettenmaier, D. P. and Kabat, P.: Hydrologic impacts of climate change on the Nile River Basin: Implications of the 2007 IPCC scenarios, Clim. Change, 100(3), 433–461, doi:10.1007/s10584-009-9693-0, 2010.

Blume, H.-P., Brümmer, G. W., Horn, R., Kandeler, E., Kögel-Knabner, I., Kretzschmar, R., Stahr, K. and Wilke, B.-M.: Lehrbuch der Bodenkunde, Springer-Verlag, Berlin Heidelberg, doi: 10.1007/978-3-662-49960-3., 2010.

Bohn, T. J. and Vivoni, E. R.: Process-based characterization of evapotranspiration sources over the North American monsoon region, Water Resour. Res., 52(1), 358–384, doi:10.1002/2015WR017934, 2016.

Breuer, L., Eckhardt, K. and Frede, H. G.: Plant parameter values for models in temperate climates, Ecol. Modell., 169(2-3), 237–293, doi:10.1016/S0304-3800(03)00274-6, 2003.

Chaney, N. W., Herman, J. D., Ek, M. B. and Wood, E. F.: Deriving global parameter estimates for the Noah land surface model using FLUXNET and machine learning, J. Geophys. Res., 121(22), 13,218–13,235, doi:10.1002/2016JD024821, 2016.

Clapp, R. B. and Hornberger, G. M.: Empirical equations for some soil hydraulic properties, Water Resour. Res., 14(4), 601–604, doi:10.1029/WR014i004p00601, 1978.

Cuntz, M., Mai, J., Samaniego, L., Clark, M., Wulfmeyer, V., Branch, O., Attinger, S. and Thober, S.: The impact of standard and hard-coded parameters on the hydrologic fluxes in the Noah-MP land surface model, J. Geophys. Res., 121(18), 10,676–10,700, doi:10.1002/2016JD025097, 2016.

Davin, E. L., Stöckli, R., Jaeger, E. B., Levis, S. and Seneviratne, S. I.: COSMO-CLM2: A new version of the COSMO-CLM model coupled to the Community Land Model, Clim. Dyn., 37(9-10), 1889–1907, doi:10.1007/s00382-011-1019-z, 2011.

Davy, R., Esau, I., Chernokulsky, A., Stephen, O. and Zilitinkevich, S.: Diurnal asymmetry to the observed global warming, Int. J. Climatol., 31(1), doi:10.1002/joc.4688, 2017.

Van Dijk, A. I. J. M., Gash, J. H., Van Gorsel, E., Blanken, P. D., Cescatti, A., Emmel, C., Gielen, B., Harman, I. N., Kiely, G., Merbold, L., Montagnani, L., Moors, E., Sottocornola, M., Varlagin, A., Williams, C. A. and Wohlfahrt, G.: Rainfall interception and the coupled surface water and energy balance, Agric. For. Meteorol., 214-215, 402–415, doi:10.1016/j.agrformet.2015.09.006, 2015.

Döll, P. and Fiedler, K.: Global-scale modeling of groundwater recharge, Hydrol. Earth Syst. Sci., 12(3), 863–885, doi:10.5194/hess-12-863-2008, 2008.

Döll, P., Trautmann, T., Gerten, D., Muller-Schmied, H., Ostberg, S., Saeed, F. and Schleussner, C.-F.: Risks for the global freshwater system at 1.5 °C and 2 °C global warming, Environ. Res. Lett., 13(4), 23, doi:10.1088/1748-9326/aab792, 2018.

Fatichi, S. and Pappas, C.: Constrained variability of modeled T:ET ratio across biomes, Geophys. Res. Lett., 44(13), 6795–6803, doi:10.1002/2017GL074041, 2017.

Gerrits, A. M. J., Savenije, H. H. G., Veling, E. J. M. and Pfister, L.: Analytical derivation of the Budyko curve based on rainfall characteristics and a simple evaporation model, Water Resour. Res., 25(4), W04403, doi:10.1029/2008WR007308, 2009.

Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W. and Sitch, S.: Terrestrial vegetation and water balance - Hydrological evaluation of a dynamic global vegetation model, J. Hydrol., 286(1-4), 249–270, doi:10.1016/j.jhydrol.2003.09.029, 2004.

Gosling, S. N. and Arnell, N. W.: Simulating current global river runoff with a global hydrological model: Model revisions, validation, and sensitivity analysis, Hydrol. Process., 25(7), 1129–1145, doi:10.1002/hyp.7727, 2011.

Gosling, S. N. and Arnell, N. W.: A global assessment of the impact of climate change on water scarcity, Clim. Change, 134(3), 371–385, doi:10.1007/s10584-013-0853-x, 2016.

Gosling, S. N., Zaherpour, J., Mount, N. J., Hattermann, F. F., Dankers, R., Arheimer, B., Breuer, L., Ding, J., Haddeland, I., Kumar, R., Kundu, D., Liu, J., van Griensven, A., Veldkamp, T. I. E., Vetter, T., Wang, X. and Zhang, X.: A comparison of changes in river runoff from multiple global and catchment-scale hydrological models under global warming scenarios of 1 C, 2 C and 3 C, Clim. Change, 141(3), 577–595, doi:10.1007/s10584-016-1773-3, 2017.

De Groen, M. M. and Savenije, H. H. G.: A monthly interception equation based on the statistical characteristics of daily rainfall., 2006.

Guo, D., Westra, S. and Maier, H. R.: An R package for modelling actual, potential and reference evapotranspiration, Environ. Model. Softw., 78, 216–224, doi:10.1016/j.envsoft.2015.12.019, 2016.

Hartmann, A., Lange, J., Weiler, M., Arbel, Y. and Greenbaum, N.: A new approach to model the spatial and temporal variability of recharge to karst aquifers, Hydrol. Earth Syst. Sci., 16(7), 2219–2231, doi:10.5194/hess-16-2219-2012, 2012.

Hartmann, A., Wagener, T., Rimmer, A., Lange, J., Brielmann, H. and Weiler, M.: Testing the realism of model structures to identify karst system processes using water quality and quantity signatures, Water Resour. Res., 49(6), 3345–3358, doi:10.1002/wrcr.20229, 2013.

Hartmann, A., Gleeson, T., Rosolem, R., Pianosi, F., Wada, Y. and Wagener, T.: A large-scale simulation model to assess karstic groundwater recharge over Europe and the Mediterranean, Geosci. Model Dev., 8(6), 1729–1746, doi:10.5194/gmd-8-1729-2015, 2015.

Hartmann, A., Gleeson, T., Wada, Y. and Wagener, T.: Enhanced groundwater recharge rates and altered recharge sensitivity to climate variability through subsurface heterogeneity, Proc. Natl. Acad. Sci., 114(11), 2842–2847, doi:10.1073/pnas.1614941114, 2017.

Haughton, N., Abramowitz, G. and Pitman, A. J.: On the predictability of land surface fluxes from meteorological variables, Geosci. Model Dev., 11(1), 195–212, doi:10.5194/gmd-11-195-2018, 2018.

Hogue, T. S., Bastidas, L. A., Gupta, H. V. and Sorooshian, S.: Evaluating model performance and parameter behavior for varying levels of land surface model complexity, Water Resour. Res., 42(8), W08430, doi:10.1029/2005WR004440, 2006.

Hong, E.-M., Pachepsky, Y. A., Whelan, G. and Nicholson, T.: Simpler models in environmental studies and predictions, Crit. Rev. Environ. Sci. Technol., 47(18), 1669–1712, doi:10.1080/10643389.2017.1393264, 2017.

Van den Hoof, C., Vidale, P. L., Verhoef, A. and Vincke, C.: Improved evaporative flux partitioning and carbon flux in the land surface model JULES: Impact on the simulation of land surface processes in temperate Europe, Agric. For. Meteorol., 181, 108–124, doi:10.1016/j.agrformet.2013.07.011, 2013.

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by V. B. and P. M. M. Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, Cambridge University Press., 2013.

IPCC: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Part B: Regional Aspects, edited by V. R. Barros, C. B. Field, D. J. Dokken, M. D. Mastrandrea, K. J. Mach, T. E. Bilir, E. S. Kissel, M. Chatterjee, K. L. Ebi, Y. O. Estrada, R. C. Genova, B. Girma, A. N. Levy, S. MacCracken, P. R. Mastrandrea, and L. L. White, Cambridge University Press., 2014.

Jarvis, P. G.: The Interpretation of the Variations in Leaf Water Potential and Stomatal Conductance Found in Canopies in the Field, Phil. Trans. R. Soc. Lond. B., 273(927), 593–610, doi:10.1098/rstb.1976.0035, 1976.

Kavetski, D. and Clark, M. P.: Numerical troubles in conceptual hydrology: Approximations, absurdities and impact on hypothesis testing, Hydrol. Process., 25(4), 661–670, doi:10.1002/hyp.7899, 2011.

Kelliher, F. M., Whitehead, D. and Pollock, D. S.: Rainfall interception by trees and slash in a young Pinus radiata D. Don stand, J. Hydrol., 131(1-4), 187–204, doi:10.1016/0022-1694(92)90217-J, 1992.

Kergoat, L.: A model for hydrological equilibrium of leaf area index on a global scale, J. Hydrol., 212-213, 268–286, doi:10.1016/S0022-1694(98)00211-X, 1998.

Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, Water Resour. Res., 42(3), W03S04, doi:10.1029/2005WR004362, 2006.

Körner, C.: Leaf diffusive conductances in the major vegetation types of the globe, in Ecophysiology of photosynthesis, edited by E. D. Schulze and M. M. Caldwell, pp. 463–490, Springer Study Edition, vol 100, Springer, Berlin, Heidelberg., 1995.

Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F. and Richardson, A. D.: Constraining a global ecosystem model with multi-site eddy-covariance data, Biogeosciences, 9(10), 3757–3776, doi:10.5194/bg-9-3757-2012, 2012.

Lenhard, J. and Winsberg, E.: Holism, entrenchment, and the future of climate model pluralism, Stud. Hist. Philos. Sci. Part B - Stud. Hist. Philos. Mod. Phys., 41(3), 253–262, doi:10.1016/j.shpsb.2010.07.001, 2010.

Lewis, S. C. and Karoly, D. J.: Evaluation of historical diurnal temperature range trends in CMIP5 models, J. Clim., 26(22), 9077–9089, doi:10.1175/JCLI-D-13-00032.1, 2013.

Liang, X., Lettenmaier, D. P., Wood, E. F. and Burges, S. J.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models, J. Geophys. Res., 99(D7), 14415–14428, doi:10.1029/94JD00483, 1994.

Lu, Y., Liu, S., Weng, L., Wang, L., Li, Z. and Xu, L.: Fractal analysis of cracking in a clayey soil under freeze-thaw cycles, Eng. Geol., 208, 93–99, doi:10.1016/j.enggeo.2016.04.023, 2016.

Medlyn, B. E., Robinson, A. P., Clement, R. and McMurtrie, R. E.: On the validation of models of forest CO2 exchange using eddy covariance data: some perils and pitfalls, Tree Physiol., 25, 839–857, doi:10.1093/treephys/25.7.839, 2005.

Mendoza, P. A., Clark, M. P., Barlage, M., Rajagopalan, B., Samaniego, L., Abramowitz, G. and Gupta, H.: Are we unnecessarily constraining the agility of complex process-based models?, Water Resour. Res., 51(1), 716–728, doi:10.1002/2014WR015820, 2015.

Miralles, D. G., Jiménez, C., Jung, M., Michel, D., Ershadi, A., McCabe, M. F., Hirschi, M., Martens, B., Dolman, A. J., Fisher, J. B., Mu, Q., Seneviratne, S. I., Wood, E. F. and Fernaìndez-Prieto, D.: The WACMOS-ET project - Part 2: Evaluation of global terrestrial evaporation data sets, Hydrol. Earth Syst. Sci., 20(2), 823–842, doi:doi.org/10.5194/hess-20-823-2016, 2016.

Mu, Q., Zhao, M. and Running, S. W.: Improvements to a MODIS global terrestrial evapotranspiration algorithm, Remote Sens. Environ., 115(8), 1781–1800, doi:10.1016/j.rse.2011.02.019, 2011.

Murray, S. J.: Trends in 20th century global rainfall interception as simulated by a dynamic global vegetation model: Implications for global water resources, Ecohydrology, 7(1), 102–114, doi:10.1002/eco.1325, 2014.

Muzylo, A., Llorens, P., Valente, F., Keizer, J. J., Domingo, F. and Gash, J. H. C.: A review of rainfall interception modelling, J. Hydrol., 370(1-4), 191–206, doi:10.1016/j.jhydrol.2009.02.058, 2009.

Pearce, A. J., Rowe, L. K. and Stewart, J. B.: Night time, wet canopy evaporation rates and the water balance of an evergreen mixed forest, Water Resour. Res., 16(5), 955–959, doi:10.1029/WR016i005p00955, 1980.

Pereira, L. S., Allen, R. G., Smith, M. and Raes, D.: Crop evapotranspiration estimation with FAO56: Past and future, Agric. Water Manag., 147, 4–20, doi:10.1016/j.agwat.2014.07.031, 2015.

Reusser, D. E. and Zehe, E.: Inferring model structural deficits by analyzing temporal dynamics of model performance and parameter sensitivity, Water Resour. Res., 47(7), W07550, doi:10.1029/2010WR009946, 2011.

Rosero, E., Yang, Z. L., Wagener, T., Gulden, L. E., Yatheendradas, S. and Niu, G.-Y.: Quantifying parameter sensitivity, interaction, and transferability in hydrologically enhanced versions of the Noah land surface model over transition zones during the warm season, J. Geophys. Res. Atmos., 115(D3), D03106, doi:10.1029/2009JD012035, 2010.

Rosolem, R., Gupta, H. V., Shuttleworth, W. J., Zeng, X. and De Gonçalves, L. G. G.: A fully multiple-criteria implementation of the Sobol' method for parameter sensitivity analysis, J. Geophys. Res., 117(D7), D07103, doi:10.1029/2011JD016355, 2012.

Running, S., Mu, Q. and Zhao, M.: MOD16A2 MODIS/Terra Net Evapotranspiration 8-Day L4 Global 500m SIN Grid V006, [Data set], , doi:10.5067/MODIS/MOD16A2.00, 2017.

Savenije, H. H. G.: Determination of evaporation from a catchment water balance at a monthly time scale, Hydrol. Earth Syst. Sci., 1(1), 93–100, doi:10.5194/hess-1-93-1997, 1997.

Scanlon, B. R., Keese, K. E., Flint, A. L., Flint, L. E., Gaye, C. B., Edmunds, W. M. and Simmers, I.: Global synthesis of groundwater recharge in semiarid and arid regions, Hydrol. Prpcesses, 20(15), 3335–3370, doi:doi.org/10.1002/hyp.6335, 2006.

Sellers, P. J., Randall, D. A., Collatz, G. J., Berry, J. A., Field, C. B., Dazlich, D. A., Zhang, C., Collelo, G. D. and Bounoua, L.: A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: Model formulation, J. Clim., 9(4), 676–705, doi:10.1175/1520-0442(1996)009<0676:ARLSPF>2.0.CO;2, 1996.

Shuttleworth, W. J.: Terrestrial Hydrometeorology, John Wiley & Sons, Ltd, Chichester, UK., 2012.

Sperna Weiland, F. C., Van Beek, L. P. H., Kwadijk, J. C. J. and Bierkens, M. F. P.: Global patterns of change in discharge regimes for 2100, Hydrol. Earth Syst. Sci., 16(4), 1047–1062, doi:10.5194/hess-16-1047-2012, 2012.

Stewart, J. B.: Modelling surface conductance of pine forest, Agric. For. Meteorol., 43(1), 19–35, doi:10.1016/0168-1923(88)90003-2, 1988.

Sugita, M. and Brutsaert, W.: Daily evaporation over a region from lower boundary layer profiles measured with radiosondes, Water Resour. Res., 27(5), 747–752, doi:10.1029/90WR02706, 1991.

Van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J. F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J. and Rose, S. K.: The representative concentration pathways: An overview, Clim. Change, 109, 5–31, doi:10.1007/s10584-011-0148-z, 2011.

Wada, Y., Van Beek, L. P. H. and Bierkens, M. F. P.: Nonsustainable groundwater sustaining irrigation: A global assessment, Water Resour. Res., 48(6), W00L06, doi:10.1029/2011WR010562, 2012.

Walter, I. A., Allen, R. G., Elliott, R., Itenfisu, D., Brown, P., Jensen, M. E., Mecham, B., Howell, T. A., Snyder, R., Eching, S., Spofford, T., Hattendorf, M., Martin, D., Cuenca, R. H. and WrightIvan, J. L.: The ASCE standardized reference evapotranspiration equation, ASCE-EWRI Task Committee Report, Environmental and Water Resources Institute of the American Society of Civil Engineers., 2005.

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O. and Schewe, J.: The Inter-Sectoral Impact Model Intercomparison Project (ISI–MIP): Project framework, Proc. Natl. Acad. Sci., 111(9), 3228–3232, doi:10.1073/pnas.1312330110, 2014.

Ye, H., Fetzer, E. J., Behrangi, A., Wong, S., Lambrigtsen, B. H., Wang, C. Y., Cohen, J. and Gamelin, B. L.: Increasing daily precipitation intensity associated with warmer air temperatures over northern Eurasia, J. Clim., 29(2), 623–636, doi:10.1175/JCLI-D-14-00771.1, 2016.

Young, P., Parkinson, S. and Lees, M.: Simplicity out of complexity in environmental modelling: Occam's razor revisited, J. Appl. Stat., 23(2-3), 165–210, doi:10.1080/02664769624206, 1996.

Zhao, Y., Ciais, P., Peylin, P., Viovy, N., Longdoz, B., Bonnefond, J. M., Rambal, S., Klumpp, K., Olioso, A., Cellier, P., Maignan, F., Eglin, T. and Calvet, J. C.: How errors on meteorological variables impact simulated ecosystem fluxes: A case study for six French sites, Biogeosciences, 9(7), 2537–2564, doi:10.5194/bg-9-2537-2012, 2012.

V2Karst V1.0: A parsimonious large-scale integrated vegetation-recharge model to simulate the impact of climate and land cover change in karst regions

Fanny Sarrazin[1], Andreas Hartmann[1,2], Francesca Pianosi[1], Rafael Rosolem[1,3], Thorsten Wagener[1,3]

[1] Department of Civil Engineering, University of Bristol, Bristol, BS8 1TR, UK
[2] Institute of Earth and Environmental Sciences, University of Freiburg, Freiburg, D-79098, Germany
[3] Cabot Institute, University of Bristol, Bristol, BS8 1UJ, UK

Correspondence to: Fanny Sarrazin (fanny.sarrazin@bristol.ac.uk)

Abstract. Karst aquifers are an important source of drinking water in many regions of the world. Karst areas are highly permeable and produce large amounts of groundwater recharge, while surface runoff is often negligible. As a result, recharge in these systems may have a different sensitivity to climate and land cover changes compared to other less permeable systems. However, little ~~effort has been directed toward assessing~~is known about the combined impact of climate and land cover changes in karst areas at large-scales. In particular, the representation of land cover and its controls on evapotranspiration has been very limited in previous karst hydrological models. In this study, we address this gap (1) by introducing the first large-scale hydrological model including an explicit representation of both karst and land cover properties, and (2) by ~~analysing~~ providing an in-depth analysis of the model's recharge production behaviour. To achieve these ~~points~~aims, we ~~first improve the~~replace the empirical approach to evapotranspiration estimation of a previous large-scale karst recharge model (VarKarst) with an explicit, mechanistic and parsimonious approach in the new model (V2Karst V1.0) ~~includes a parsimonious representation of relevant ET processes for climate and land cover change impact studies~~. We demonstrate the plausibility of V2Karst simulations at four carbonate rock FLUXNET sites by assessing the model's ability to reproduce observed evapotranspiration and soil moisture patterns and~~using soft rules and global sensitivity analysis~~ by showing that the controlling modelled processes are in line with expectations. Additional virtual experiments with synthetic input data systematically explore the sensitivities of recharge to precipitation characteristics (overall amount and temporal distribution) and land cover properties. This approach confirms that these sensitivities agree with expectations and provides first insights into potential impacts of future change. V2Karst is the first model that enables the study of the joint impacts of large-scale land cover and climate changes on groundwater recharge in karst regions. ~~Then, we use virtual experiments with synthetic data to assess the sensitivity of simulated recharge to precipitation characteristics and land cover. Results reveal how both vegetation and soil parameters control the model behaviour, and they suggest that simulated recharge is sensitive to both precipitation (overall amount and temporal distribution) and land cover.~~ ~~Large-scale assessment of future karst groundwater recharge should therefore consider~~

~~the combined impact of changes in land cover and precipitation properties, if it is to produce realistic projections of future change impacts.~~

1. Introduction

Carbonate rocks, from which karst systems typically develop, are estimated to cover 10-15% of the world ~~continental areas~~ (Ford and Williams, 2007). More importantly, karst aquifers are a considerable source of drinking water for almost a quarter of the world's population (Ford and Williams, 2007) and ~~have~~ play a critical role in sustaining food production because most karst areas ~~present~~ contain some form of agricultural activity (Coxon, 2011). In Europe, carbonate rock areas cover 14-29% of the land area, and some European countries such as Austria and Slovenia derive up to 50% of their total water supply from karst aquifers (Chen et al., 2017; COST, 1995).

Karst systems are characterised by a high spatial variability of bedrock and soil permeability due to the presence of preferential flow pathways (Hartmann et al., 2014). The soluble carbonate bedrock is structured by large dissolution fissures or conduits (Williams, 1983, 2008) and the typically clayey soil often contains cracks (Blume et al., 2010; Lu et al., 2016) where infiltrating water concentrates. Therefore, a large part of the groundwater recharge occurs as concentrated and fast flow in large apertures ~~and~~ while the other part moves as diffuse and slow flow through the matrix (Hartmann and Baker, 2017). Preferential flow pathways are particularly developed in karst, but they are also widely found in many other systems, due to root and organism activities, discontinuous subsurface layers, surface depressions, soil desiccation, tectonic processes and physical and chemical weathering (Beven and Germann, 2013; Hendrickx and Flury, 2001; Uhlenbrook, 2006)

Preferential infiltration is typically triggered when thresholds in rain intensity and soil moisture levels are exceeded (Rahman and Rosolem, 2017; Tritz et al., 2011). When activated, preferential infiltration pathways may enhance groundwater recharge while limiting surface runoff (e.g. Bargués Tobella et al., 2014). In karst systems, permeability is often so high that surface runoff is negligible, and virtually all precipitation infiltrates (Contreras et al., 2008; Fleury et al., 2007; Hartmann et al., 2014). Furthermore, preferential infiltration pathways can affect the temporal dynamics of recharge. For instance, Cuthberth et al. (2013) showed that macro-pores in the soil can generate quick responses in the water table, and Arbel et al. (2010) observed that dripping rates in a karst cave can fluctuate following precipitation inter-seasonal and intra-seasonal variations.

Changes in weather patterns ~~(e.g. due to climate change)~~, ~~and~~ specifically in the intensity and frequency of precipitation, may alter the activation of preferential flow pathways. ~~From previous studies in non-karst areas we can learn that changes in the intensity and frequency of precipitation events have an impact on the water yield. For instance, using an analytical framework and synthetic experiments, (Porporato et al., 2004) established a dependency between the soil water balance and both the frequency and intensity of precipitation events, while (Jothityangkoon and Sivapalan, 2009) determined different theoretical hydrological regimes based on the intensity and frequency of the precipitation input over Australian catchments. A modelling study by (Weiß and Alcamo, 2011) showed that for a given change in the total precipitation amount, a change in the intensity of precipitation events have a larger impact on water availability than a change in the number of wet days over European river basins. Regarding groundwater recharge, observation records are scarce but data indicate a sensitivity to extreme rainfall in a semi-arid tropical region (Taylor et al., 2013;) and in a seasonally humid tropical region (Owor et al.,~~

~~2009).~~ In karst areas, ~~few~~ several modelling studies showed that groundwater recharge (Hartmann et al., 2012; Loáiciga et al., 2000), spring discharge (Hao et al., 2006), and streamflow (Samuels et al., 2010) respond to changes in climate. However, to our knowledge, only the study by Hartmann et al. (2017) quantified the sensitivity of simulated karst groundwater recharge to specific precipitation characteristics, namely the mean precipitation and the intensity of daily heavy precipitation events. This modelling study was conducted across carbonate rock areas in Europe, the Middle East and Northern Africa ~~,~~ for different climate change projections. The study ~~authors~~ results suggested~~s~~ that, due to the presence of preferential flow pathways, recharge in karst systems tends to exhibit a higher sensitivity to changes in mean precipitation and ~~to~~ in the intensity of heavy precipitation events in dry climates, and a lower sensitivity in wet climates compared to non-karst systems. Hartmann et al. (2017) also showed that the intensity of heavy precipitation can have both a positive or a negative effect on recharge for both karst and non-karst systems~~,~~. However, other observational studies in non-karst areas associate increases in extreme rainfall with higher recharge amount, e.g. in a semi-arid tropical region (Taylor et al., 2013) and in a seasonally humid tropical region (Owor et al., 2009). The discrepancy among these results might be explained by the fact that Hartmann et al. (2017) only tested recharge sensitivity against precipitation properties, while ignoring their interactions with other meteorological variables such as temperature or humidity.

In addition to climate change, land cover/use change ~~could also~~is also expected to have a major impact on hydrological processes in the future (DeFries and Eshleman, 2004; Vörösmarty, 2002). Changes in land cover/use can impact the partitioning between green water (water that can be lost through evapotranspiration ~~losses~~) and blue water (water potentially available for human activities, namely groundwater recharge and runoff) (Falkenmark and Rockström, 2006). Green water tends to be higher for forested areas than for shorter vegetation (e.g. Brown et al., 2005), which has also been ~~found in few local studies in~~confirmed for karst areas (Ford and Williams, 2007; Williams, 1993). Significant land cover/use changes are expected to occur in the future, including in European and Mediterranean karst areas. These ~~are~~ will partly be ~~due to~~caused by modifications in socio-economic and technological factors, such as changes to food and wood demand or changes in agricultural management practices that could enhance agricultural yields (see e.g. Holman et al., 2017 for a European assessment; Hurtt et al., 2011 for a global assessment). Future vegetation will also be impacted by other changing environmental conditions such as ~~modifications~~ increases in atmospheric CO_2 leading to differences in plant behaviour and leaf area index, as well as to differences in climate and weather patterns, which in turn change the frequencies ~~, nitrate deposition and climate, and~~of natural disturbances such as wildfires, storms or bark beetle infestations ~~could also cause changes in land cover and in vegetation characteristics (e.g. leaf area index)~~ (Seidl et al., 2014; Zhu et al., 2016).

The above review of the literature reveals that changes in climate characteristics (e.g. precipitation intensity and frequency) and in land cover properties ~~are~~ can be expected to have significant ~~combined~~ impacts on ~~karst~~ the hydrology of karst regions. Yet, the ~~impact of preferential pathways on the partitioning between green and blue water, and the~~ joint effect of ~~climate and~~

~~land cover change~~changes in these boundary conditions has not been studied systematically, while such assessment is needed at large-scales to inform water resources management plans (Archfield et al., 2015). ~~for those areas, thus creating uncertainties in future groundwater estimates and therefore~~. ~~Determining how sensitive groundwater recharge is to climate and land cover may change in the presence of preferential pathways, and therefore to what extent findings obtained for non-karst areas may be extrapolated to karst ones, is also essential to improve our understanding of future groundwater recharge at large-scales and ultimately to improve~~ In this study we introduce a novel large-scale hydrological model that includes explicit representation of both karst and vegetation properties, and systematically explore the sensitivity of its ~~its simulated~~ groundwater recharge predictions to meteorology and vegetation ~~inputs~~properties. Our model ~~builds~~ extends~~on~~ an existing karst hydrology model, called VarKarst, which was recently developed for large-scale applications and ~~demonstrated~~ tested over European and Mediterranean carbonate rock areas (Hartmann et al., 2015). However, VarKarst only contains~~has~~ a ~~very~~ simplistic and empirical representation of evapotranspiration processes that ~~and~~ does not include land cover properties explicitly, which~~, up to now,~~ prevented its application in land cover change impact studies.

~~The present~~Our study has two objectives~~ that help us to overcome the previous limitation~~. First, we ~~aim to~~ add an explicit representation of land cover properties to VarKarst by improving its evapotranspiration (ET) estimation. While we seek to keep the model structure parsimonious, we want the new version of the model~~, called (~~ V2Karst ~~(~~V1.0)~~,~~ to be appropriate for assessing hydrological impacts of combined land cover and climate change ~~impact studies to advance~~at large-scale ~~hydrologic modelling~~. ~~-~~The model should ultimately simulate~~with respect to~~ ~~seasonal and annual~~ groundwater recharge, which is the amount of renewable groundwater available to human consumption and ecosystems, at both seasonal and annual time scales (e.g. Döll and Fiedler, 2008; Scanlon et al., 2006; Wada et al., 2012). Second, we ~~aim to~~ ~~We~~ test the plausibility of the V2Karst model behaviour by comparing its predictions against observations available at carbonate rock FLUXNET sites, and by analysing the sensitivity of simulated recharge using both measured and synthetic data to force the model. In particular, the use of synthetic data in virtual experiments ~~dominant controls of simulated recharge. Second, we aim to understand the sensitivity of simulated groundwater recharge with V2Karst to changes in the vegetation characteristics and climate. We use a set of virtual experiments that~~ allows us to control variations in climate and vegetation inputs, so that we can ~~better explore~~isolate their individual and combined effects on ~~model outputs~~simulated recharge~~.~~ and overcome~~s~~ some of the ~~previous issues of isolating the impact of individual variables as~~issues found by Hartmann et al. (2017) as discussed above.

2 New version of VarKarst with explicit representation of land cover properties (V2Karst)

~~In this section, we first introduce our rationale to explicitly represent land cover properties into VarKarst (Sec. 2.1), we then briefly describe the previous ET component of VarKarst (Sect. 2.2) and we finally present the new V2Karst model (Sect. 2.3).~~

2.1 Rationale ~~to~~ for our approach to ~~explicitly represent~~ land cover ~~properties~~ representation in~~to~~ VarKarst

The new version of the VarKarst model should be appropriate to assess the impact of climate and land cover change on karst groundwater recharge. It should also consider the range of challenges related to modelling ET at large-scales, namely a lack of ET observations to compare with model predictions~~,~~ and a lack of observations of vegetation properties (e.g. rooting depth, stomatal resistance, canopy interception storage capacity)~~, and uncertainty in large-scale forcing weather variables (specifically air temperature, net radiation, humidity and wind speed).~~ to constrain the model parameters (See section S1 in our Supplementary material). ~~Further details on these three challenges are reported in Sect. S1 of our supplementary material. According to that, w~~We ~~hence~~ define the three following criteria to develop an ~~represent ET in the~~ enhanced version of the VarKarst model with explicit representation of land cover properties:

1. **The new model should assess ~~separately all~~ the three main ET components (bare soil evaporation in presence of sparse canopy, transpiration and evaporation from canopy interception) separately and explicitly.** In fact, these fluxes exhibit different dynamics and sensitivity to environmental conditions; therefore they are likely to respond differently to climate and land cover changes (Gerrits, 2010; Maxwell and Condon, 2016; Savenije, 2004; Wang and Dickinson, 2012).
2. **The model should therefore use a Penman-Monteith formulation ~~for~~ to estimate the potential evapotranspiration (PET)** (Monteith, 1965), so that it can~~to~~ separate the effects of climate and land cover ~~and assess specific rates for~~on each of the ~~different~~ ET components. In fact, empirical PET formulations such as the Priestley-Taylor equation (Priestley and Taylor, 1972) do not allow for such separations as they do not ~~represent~~ explicitly include land cover properties.
3. **All processes should be represented parsimoniously** in accordance with the modelling philosophy underpinning the first version of VarKarst (Hartmann et al., 2015). This criterion aims to avoid over-parameterisation given the limited amount of available information to constrain and test model simulations in particular at large-scales (Abramowitz et al., 2008; Beven and Cloke, 2012; Haughton et al., 2018; Hong et al., 2017; IPCC, 2013, Chapter 9, pp. 790-791; Young et al., 1996). In particular, parameters that account for physical properties of the system (e.g. soil and vegetation properties) are commonly ~~read~~taken from look-up tables but their physical meaning have been put into question and they may actually not be commensurate with field measurements as discussed in Hogue et al. (2006) and in Rosero et al. (2010). Therefore, it has been suggested that even these 'physical' parameters should be calibrated so to optimise the model performance (Chaney et al., 2016; Rosolem et al., 2013). Importantly, parsimony~~,~~ limits the computational time for model simulations and allows for assessing the impact of modelling choices and the uncertainty and sensitivity of model output using Monte-Carlo simulation (Hong et al., 2017; Young et al., 1996).

~~We review the different approaches currently used to explicitly represent explicitly land cover properties in existing hydrological models of karst areas and in existing large-scale models, to assess their consistency with the three above-mentioned criteria and to determine whether we could directly adopt some of these ET representations for VarKarst. We report~~

~~a summary of our findings in the following paragraphs, while more details on the processes and parameterisations of the ET components of other large-scale models are reported in Tables (A1-A3), and a detailed list of all parameters involved in the representation of ET in large-scale models can be found in Section S2 of our supplementary material.~~

With respect to ~~existing~~ previous modelling studies of karst systems, to our knowledge, only four have used models that explicitly include ~~ET and~~ land cover pro~~cesses~~perties ~~were applied in karst studies~~, all of which were applied to the local scale ~~studies where~~with detailed on-site information ~~was available~~. Three of these studies (Canora et al., 2008; Doummar et al., 2012; Zhang et al., 2011) used generic hydrological models that were not specifically developed for karst areas but ~~they are distributed~~ included enough ~~models that simply utilised the~~ flexibility in their spatially-distributed parameters to represent the variability in soil and bedrock properties. ~~These models are~~ The ~~heavily~~ large number of parameters in these models ~~ised, which~~ hampers their application at large-scales~~,~~ and does not comply with criterion 3 (parsimony). The fourth study ~~introduced in~~ (Sauter, 1992) used a ~~is~~ lumped model that ~~and~~ is much more parsimonious ~~than the three other models~~but ~~. However, the model~~ does not represent soil evaporation~~,~~ and uses empirical PET equations, which does not allow to separate the effect of climate and land cover (~~disagreement with~~ criteria 1 and 2). ~~Moreover, the model of (Sauter, 1992) has a rather sophisticated interception routine, which includes both canopy and trunk interception.~~

As for large-scale models, we can identify two main types: hydrological models, which focus on the assessment of hydrological fluxes, and land surface models, which also evaluate many other fluxes ~~(~~such as sensible heat, latent heat, ground heat flux, radiation and carbon fluxes (for a review, see Bierkens, 2015). Land surface models do not usually comply with criterion~~a~~ 3 because they have many parameters, including a number of empirical parameters that are difficult to constrain (Mendoza et al., 2015). Moreover, it has been shown that land surface models could be simplified when the objective is to assess hydrological fluxes only. For instance, Cuntz et al. (2016) demonstrated that a large number of parameters of the Noah land surface model are non-influential or have very little influence on simulated runoff. In contrast, hydrological models focus on the representation of hydrological processes and include ~~fare~~far fewer parameters. However, ~~Regarding the representation of ET in existing large-scale models,~~ our literature review (summarised in Tables A1-A3) showed that we cannot directly adopt any of their ET representation into VarKarst. In fact, as shown in Tables (A1-A3), the most parsimonious models (WBM~~, LaD and~~ WaterGap and mHM) neglect some ET components and/or use empirical PET equations, which contradicts criteria 1 and 2, while models that comply with criteria 1 and 2 (PCR-GLOBWB, ~~ISBA and~~ VIC and the model of Kergoat (1998)) use heavily parameterised schemes, such as a Jarvis type parameterisation of surface resistance (Jarvis, 1976; Stewart, 1988) and therefore do not satisfy criterion 3 (parsimony). Moreover, we found that large-scale models include empirical schemes with no clear origin, such as the reference crop formulation used in the PCR-GLOBWB model for PET calculation or the interception model used in LPJ and in the model of Kergoat (1998). Importantly, our review revealed the tremendous variability of approaches used in large-scale models ~~when it comes to representing~~to represent ET processes. A detailed list of all parameters involved in the representation of ET the models of Table A1-A3 can be found in Section S2 of our Supplementary material. Consequently, no clear indication emerged regarding a ‘best way’ to parameterise the different ET processes at large-scales, which leaves us with a large range of different formulations to choose from to implement an explicit representation of land cover processes into VarKarst.

~~The next sections provide more details on the specific assumptions and choices made to develop the new ET component for the VarKarst model, which satisfies the three criteria defined in this section and utilises some of the schemes from other large-scale models.~~

2.2 Previous representation of ET processes in VarKarst

VarKarst (Hartmann et al., 2015) ~~[Hartmann et al., 2015]~~ is currently the only karst recharge model developed for large-scale applications. It is a conceptual semi-distributed model that simulates karst potential recharge (Fig. 1.a). VarKarst includes two horizontal subsurface layers, a top layer called ‘soil’ and a deeper layer called ‘epikarst’. The soil layer corresponds to the layer from which ET can occur. The epikarst layer corresponds to the uppermost layer of weathered carbonate rocks where it is assumed that water cannot be lost through ET. ~~Groundwater recharge predicted by VarKarst includes both the diffuse and concentration fractions,~~ VarKarst represents karst processes because for each model grid cell, the water balance is evaluated separately over a number of vertical compartments with varying soil and epikarst properties. Additionally, and lateral flow concentrates the infiltrating water from saturated to unsaturated compartments. Conceptually, the direct contribution of precipitation to infiltration and recharge can be associated with the diffuse fraction, while the contribution of lateral flow can be associated with the concentrated fraction. The model has been shown to be more appropriate for applications over karst areas than other large-scale hydrological models that do not represent karst processes (Hartmann et al., 2017). The ET component of the VarKarst model is very simple and does not include explicit representation of land cover properties. ET is lumped in the soil layer, is estimated from PET and reduced by a water stress factor, which is estimated as a linear function of soil moisture. The PET rate is calculated with the empirical Priestley-Taylor equation (Priestley and Taylor, 1972) using a spatially and temporally uniform value of the empirical coefficient. This approach does not allow to separate the effect of climate and land cover, since the empirical coefficient reflects both climate and vegetation characteristics simultaneously. Therefore, the ET component of VarKarst needs to be modified if the model is to be used for large-scale land cover change impact assessment.

2.3 V2Karst: the new version of VarKarst for integrated vegetation-recharge simulations over karst areas

In this section, we propose a new version of the VarKarst model, called V2Karst (~~Figure~~ Fig. 1b). In accordance with criteria 1 and 2 defined in Sect. 2.1~~, compared to VarKarst~~, the new V2Karst model includes a physically based PET equation, separates the evapotranspiration flux into three components (transpiration, bare soil evaporation and evaporation from canopy interception), and comprises three soil layers. In agreement with criteria 3 of Sect. 2.1 (parsimony), we sought to represent parsimoniously the different ET processes into VarKarst. In fact, V2Karst uses 12 parameters to represent ET and vegetation seasonality ~~(~~, ~~including~~namely ~~the~~ 11 newly introduced parameters that replace the Priestley-Taylor empirical coefficient α used in Varkarst and the soil water capacity parameters~~s~~ V_{soi} already present in VarKarst (model parameters are described in Table 1 and Fig. 1)~~)~~. This is less than other existing large-scale models that use Penman-Monteith equation and separate the three ET components, since these models have over 15 parameters in their ET component (PCR-GLOBWB, ~~ISBA and~~ VIC and model of Kergoat (1998) in Tables A1-A3). ~~Additionally, V2Karst represents parsimoniously the seasonal changes in the vegetation properties, which will allow us to analyse the importance of this process on simulated recharge.~~ We assumed homogeneous above ground vegetation properties across model compartments.

~~We note that V2Karst has a total of 15 parameters (described in Table 1 and Figure 1), including the 4 parameters of VarKarst and 11 new parameters in the new ET component, that replaces the Priestley-Taylor empirical coefficient α used in VarKarst. In agreement with criteria 3 of Sect. 2.1 (parsimony), we sought to represent parsimoniously the different ET processes into VarKarst. In fact, V2Karst uses 12 parameters to represent ET and vegetation seasonality (including the 11 newly introduced parameters and the soil water capacity parameters V_{soi} already present in VarKarst). This is less than other existing large-scale models that use Penman-Monteith equation and separate the three ET components, since these models have over 15 parameters in their ET component (PCR-GLOBWB, ISBA and VIC and model of (Kergoat, 1998) in Tables A1-A3).~~ The new model is forced by time series of precipitation P, air temperature T and net radiation R_n as VarKarst. Additionally, time series of relative humidity RH and wind speed WS are now needed for PET calculation. The V2Karst model can be run at both daily and sub-daily time step. In this study, we present simulation results obtained using a daily time step, while results from hourly simulations are reported in Sect. S8 of our Supplementary Material. We did not find significant differences between daily and hourly simulations for assessing recharge at monthly and annual time scale, which is the focus of our study. Therefore, it is reasonable to apply the model at daily time step, which significantly reduces the computational requirements.

2.3.1 Definition of soil and epikarst properties in V2Karst

The computation of water storage capacity of the entire soil column $V_{S,i}$ [mm] and of the epikarst $V_{E,i}$ [mm], and the epikarst outflow coefficient $K_{E,i}$ [d] for the ith model compartment is done as before in VarKarst:

$$
\begin{aligned}
V_{S,i} &= V_{max,S}\left(\frac{i}{n_c}\right)^a, \\
V_{E,i} &= V_{max,E}\left(\frac{i}{n_c}\right)^a, \\
K_{E,i} &= K_{max,E}\left(\frac{n_c - i + 1}{n_c}\right)^a.
\end{aligned}
\tag{1}
$$

where $V_{max,S}$ [mm] is the maximum soil storage capacity over all model compartments, $V_{max,E}$ [mm] is the maximum epikarst storage capacity, $K_{max,E}$ [d] is the maximum outflow coefficient, n_c [−] is the number of model compartments, which is set to 15 following (Hartmann et al., 2013, 2015) and a [−] is the spatial variability coefficient. A previous study showed that $V_{S,i}$, $V_{E,i}$ and $K_{E,i}$ can be determined using the same distribution coefficient a (Hartmann et al., 2013). In V2Karst, $V_{max,S}$, $V_{max,E}$ and $K_{E,i}$ are computed as a function of the average properties of the cell using the following ~~the~~ formulas:

$$V_{max,S} = \frac{V_{soi} n_c}{\sum_{i=1}^{n} \left(\frac{i}{n_c}\right)^a},$$

$$V_{max,E} = \frac{V_{epi} n_c}{\sum_{i=1}^{n} \left(\frac{i}{n_c}\right)^a}, \tag{2}$$

$$K_{max,E} = \frac{K_{epi} n_c}{\sum_{i=1}^{n} \left(\frac{i}{n_c}\right)^a},$$

where V_{soi} [mm] is the mean soil storage capacity, V_{epi} [mm] is the mean epikarst storage capacity and K_{epi} [mm] is the mean epikarst outflow coefficient. We note that the definition of the three parameters V_{soi}, V_{epi} and K_{epi} is revised compared to VarKarst.

As in VarKarst, we neglect ET from the epikarst. Several studies showed that in presence of shallow soil and dry climate, plants can take up water in the weathered bedrock where soil pockets can sustain roots development (Schwinning, 2010). However, given the uncertainty in soil depth for large-scale applications, V2Kast does not allow ET from the epikarst to avoid over-parameterisation. Therefore, the V2Karst soil layer must be interpreted as a conceptual layer that does not exactly correspond to the physical soil layer (layer of loose material) but is defined as the portion of the subsurface where ET losses can occur.

In V2Karst, the soil layer is further divided into a shallow top layer from which water can be lost from both evaporation and transpiration, a second middle layer where only transpiration can occur and a third deeper layer below the root zone where transpiration can only take place when the first two layers are depleted. The maximum storage capacity of the first layer is noted as V_e [mm], and the maximum storage capacity of first and second layers combined is noted as V_r [mm], which corresponds to the maximum storage capacity of the root zone. The model assumes that V_e is smaller than V_r, which is in turn smaller than the storage capacity of the deeper model compartment $V_{S,n}$.

2.3.2 Soil water balance

The soil water storage $V_{soi,i}^{j}(t)$ [mm] in the ith compartment and the jth soil layer $j = 1,2,3$ is updated at the end of each time step t as follows:

$$\begin{aligned} V_{soi,i}^{1}(t) &= V_{soi,i}^{1}(t-1) + T_f(t) + Q_{lat,i-1\to i}(t) - Es_{act,i}(t) - T_{act,i}^{1}(t) - R_{12,i}(t), \\ V_{soi,i}^{2}(t) &= V_{soi,i}^{2}(t-1) + R_{12,i}(t) - T_{act,i}^{2}(t) - R_{23,i}(t), \end{aligned} \tag{3}$$

$$V_{soi,i}^{3}(t) = V_{soi,i}^{3}(t-1) + R_{23,i}(t) - T_{act,i}^{3}(t) - R_{epi,i}(t).$$

where $T_f(t)$ [mm] is the throughfall i.e. the fraction of precipitation that ~~is not evaporated from the interception store~~reaches the ground (Eq. (8)), $Q_{lat,i-1\rightarrow i}(t)$ [mm] is the lateral flow from the $(i-1)$th to the ith model compartment (Sect. 2.3.4), $Es_{act,i}(t)$ [mm] is the actual soil evaporation (Eq. (9~~7~~)), $T_{act,i}^{j}(t)$ [mm] is the actual transpiration in the jth soil layer (Eq. (~~9-10~~11-12)), $R_{12,i}(t)$ [mm] is the downward flow from the first to the second soil layer, $R_{23,i}(t)$ [mm] is the downward flow from the second to the third soil layer and $R_{epi,i}(t)$ [mm] is the downward flow from the soil to the epikarst.

It is assumed that percolation from the unsaturated soil to the epikarst is negligible due to low permeability of the soil. This assumption seems reasonable since karst soils usually have a high clay content (Blume et al., 2010; Clapp and Hornberger, 1978). However, clayey soil typically present cracks (Lu et al., 2016), and therefore when the soil reaches saturation, preferential flow starts to occur in the soil cracks, which causes all saturation excess to quickly infiltrate to the epikarst. Just as in VarKarst, such preferential vertical flow is represented by the variable $R_{epi,i}(t)$ (used in Eq. (3)) and is set equal to the saturation excess in the (lowest) soil layer. In V2Karst, a similar approach is also used to assess the other vertical flows from one soil layer to another ($R_{12,i}(t)$ and $R_{23,i}(t)$) in Eq. (3)).

2.3.3 Evapotranspiration

We adopt the representation of sparse vegetation proposed by Bohn and Vivoni (2016) for the VIC model and referred to as ‘clumped’ vegetation scheme. Each model compartment is divided into a vegetated and a non-vegetated fraction using a canopy cover fraction coefficient $f_c(t)$ $[-]$. The uptake of soil moisture for transpiration and soil evaporation is coupled in a way that, for each model compartment, we evaluate an overall water balance over the two fractions. Using such a coupled approach facilitates the representation of the seasonal variations in vegetated and non-vegetated fractions compared to an uncoupled ‘tile’ approach, in which a separate soil moisture state is represented for vegetated and bare soil fractions. Consistently with other existing large-scale models, aerodynamic interactions between both fractions are neglected to keep the number of parameters to a minimum (Table A3).

The canopy coefficient $f_c(t)$ is estimated in V2Karst using the Beer-Lambert’s law as in Van Dijk and Bruijnzeel (2001) and Ruiz et al. (2010). This law has been originally used to separate the fraction of incident radiation (and by extension of net radiation) absorbed by the canopy from the fraction penetrating the canopy (Kergoat, 1998; Ross, 1975; Shuttleworth and Wallace, 1985). The canopy cover fraction at time t is expressed as a function of the cell average leaf area index $LAI(t)$ $[\mathrm{m}^2\ \mathrm{m}^{-2}]$ and an extinction coefficient k $[-]$, which is understood to vary across vegetation type since it accounts for leaf architecture (Ross, 1975):

$$f_c(t) = 1 - e^{-kLAI(t)}. \quad (4)$$

Notice that Eq. (4) allows to describe the seasonal variations in canopy cover fraction without introducing additional parameters in the model, given that it will simply follow the seasonal variations in LAI.

Canopy interception

The interception storage capacity over the vegetated fraction $V_{can,max}(t)$ [mm] depends (1) on the leaf area index over the vegetated fraction, which is estimated by rescaling the cell average leaf area index $LAI(t)$ by the vegetation cover fraction $f_c(\mathrm{t})$ (as in Bohn and Vivoni, 2016)~~,~~ and (2) on the canopy storage capacity per unit of leaf area index, denoted by V_{can}, which is understood to depend on the vegetation type since it accounts for leaf architecture (Gerrits, 2010). Specifically, ~~I~~it is expressed as:

$$V_{can,max}(\mathrm{t}) = V_{can}\left(\frac{\mathrm{LAI(t)}}{f_c(\mathrm{t})}\right). \qquad (5)$$

The interception storage over the vegetated fraction $I_c(t)$ [mm] is then updated at each time step as follows:

$$I_c(t) = min\left(P(t) + I_c(t-1) - \frac{Ec_{act}(t)}{f_c(t)}, V_{can,max}(t)\right). \qquad (6)$$

The actual evapotranspiration from canopy interception $Ec_{act}(t)$ $[mm]$ is computed as:~~It has been shown that a simple parameterization of daily interception can give reasonable simulation results (Gerrits, 2010; De Groen, 2002; Savenije, 1997). Following these studies, in V2Karst, interception is represented by a daily threshold model. Our formulation is as follows:~~

$$Ec_{act}(t) = f_c(t)min\left(Ec_{pot}(t), P(t) + I_c(t-1), V_{can,max}(t)\right), \qquad (\cancel{5}7)$$

where $Ec_{pot}(t)$ [mm] is the potential evaporation from canopy interception (Eq. (14~~2~~)) and~~,~~ $P(t)$ [mm] is the precipitation ~~and $V_{can,max}(t)$ [mm] is the interception storage capacity over the vegetated fraction of the cell (Eq. (6))~~. The factor $f_c(t)$ in Eq. (~~5~~7) accounts for the fact that evaporation from canopy occurs over the vegetated fraction only. ~~We note that the potential rate $Ec_{pot}(t)$ was not accounted for in the original formulation by (Gerrits, 2010; De Groen, 2002; Savenije, 1997).~~ Finally, the throughfall is calculated as:

$$T_f(t) = \max\left(P(t) - Ec_{act}(t) - f_c(t)\left(I_c(t) - I_c(t-1)\right), 0\right). \qquad (8)$$

~~The interception storage capacity over the vegetated fraction $V_{can,max}$ [mm] depends (1) on the leaf area index over the vegetated fraction, which is estimated by rescaling cell average leaf area index $LAI(t)$ using the vegetation cover fraction $f_c(\mathrm{t})$ following (Bohn and Vivoni, 2016), and (2) on the canopy storage capacity per unit of leaf area index, denoted by V_{can}, which is understood to depend on the vegetation type since it accounts for leaf architecture (Gerrits, 2010). It is expressed as:~~

$$\cancel{V_{can,max}(\mathrm{t}) = V_{can}\left(\frac{\mathrm{LAI(t)}}{f_c(\mathrm{t})}\right).} \qquad \cancel{(6)}$$

Previous studies suggest that daily simulations of interceptions can provide reasonable results (Gerrits, 2010; De Groen, 2002; Savenije, 1997). For daily simulation, the model does not account for the carry-over of interception storage from one day to the next, which means that $I_c(t)$ is set to zero at the end of each day and that all precipitation which is not evaporated from the interception store ~~reaches the ground as throughfall~~becomes throughfall as in Gerrits (2010), De Groen (2002) and Savenije (1997). This assumption can be justified by the fact that the interception process is highly dynamic at a sub-daily time scale, because the canopy can go through several wetting-drying cycles within a day (Gerrits, 2010). Therefore, ~~when evaporation from canopy interception is estimated with a~~at daily time step~~ as in V2Karst~~, the canopy layer must be interpreted as a conceptual layer, whose storage capacity does not exactly correspond to the physical storage capacity of the canopy (i.e. the amount of water that can be hold at a given time)~~,~~ but to the cumulative amount of water that can be hold by the canopy over a day (Gerrits, 2010).

Bare soil evaporation

It is assumed that soil evaporation is a faster process than transpiration consistently with general knowledge on ET processes (Wang and Dickinson, 2012). Therefore, soil moisture can be first evaporated and then transpired if some available moisture remains for plant water uptake. Soil evaporation is withdrawn for the first soil layer as a function of the potential rate and soil moisture, similar to the previous version of VarKarst:

$$Es_{act,i}(t) = \min\left(\left(1 - f_c(t)\right)Es_{pot}(t)\frac{V^1_{soi,i}(t-1)}{V^1_{S,i}}, V^1_{soi,i}(t-1) + T_f(t)\right), \qquad (\cancel{7}9)$$

where $Es_{pot}(t)$ is the potential soil evaporation (Eq. (~~12~~14)). The factor $(1 - f_c(t))$ in Eq. (~~7~~9) accounts for the fact that soil evaporation occurs from the non-vegetated fraction only and therefore the potential rate has to be weighted by the bare soil cover fraction. The right term of the equation $(V^1_{soi,i}(t-1) + T_f(t))$ is not weighted because we assume that the soil moisture is uniform over the fractions of each model compartment (we compute a unique water balance) and therefore the total moisture present in the first soil layer is available to soil evaporation because the vegetated fraction can supply moisture to the bare soil fraction.

Transpiration ~~from vegetated soil~~over the vegetated fraction

Transpiration mainly occurs in the first and second soil layers, and it switches to the third soil layer when the first two layers are depleted. The extraction of water by the roots below the root zone is e.g. documented in Penman (1950).~~ and we~~ We account for this process by representing a soil layer below the root zone, which can provide water to the root zone through capillary rise as in the ISBA model (Boone et al., 1999). In V2Karst, the rate at which transpiration occurs in the two first soil layers $T^{12}_{rate,i}(t)$ [mm] and in the third soil layer $T^3_{rate,i}(t)$ [mm] are assessed as follows:

$$T^{12}_{rate,i}(t) = \left(1 - t_{wet}(t)\right)f_c(t)T_{pot}(t)\frac{V^1_{soi,i}(t-1) + V^2_{soi,i}(t-1)}{V^1_{S,i} + V^2_{S,i}}, \qquad (\cancel{8}10)$$

$$T^3_{rate,i}(t) = \big(1 - t_{wet}(t)\big) f_c(t) T_{pot}(t) \frac{V^3_{soi,i}(t-1)}{V^3_{S,i}} f_{red}.$$

Where $T_{pot}(t)$ is the potential transpiration (Eq. (1~~4~~2)), $t_{wet}(t)$ $[-]$ is the fraction of the ~~day~~ time step with wet canopy (Eq. (~~11~~13)) and f_{red} $[-]$ is a reduction factor which accounts for the fact that moisture below the root zone is less easily accessible to the roots than moisture in the root zone (Penman, 1950), and which is expected to vary across soil type since it is linked to the soil capability to supply water to the root zone. It is assumed that transpiration occurs in the two first soil layers when $T^{12}_{rate,i}(t)$ is higher than $T^3_{rate,i}(t)$, and that transpiration is drawn from the third soil layer otherwise. The actual transpiration in the two first soil layers $T^{12}_{act,i}(t)$ [mm] and in the third soil layer $T^3_{act,i}(t)$ [mm] are therefore calculated as follows:

$$\begin{aligned}
&\text{when } T^{12}_{rate,i}(t) \geq T^3_{rate,i}(t):\\
&\begin{cases} T^{12}_{act,i}(t) = \min\Big(T^{12}_{rate,i}(t), V^1_{soi,i}(t-1) + V^2_{soi,i}(t-1) + T_f(t) - Es_{act,i}(t)\Big), \\ T^3_{act,i}(t) = 0, \end{cases}\\
&\text{when } T^{12}_{rate,i}(t) < T^3_{rate,i}(t):\\
&\begin{cases} T^{12}_{act,i}(t) = 0, \\ T^3_{act,i}(t) = \min\Big(T^3_{rate,i}(t),\ V^3_{soi,i}(t-1) + R_{23,i}(t)\Big). \end{cases}
\end{aligned}$$

(~~9~~11)

Actual transpiration in the upper two layers $T^{12}_{act,i}(t)$ is partitioned between the two soil layers within the root zone as is used in the PCR-GLOBWB model (Van Beek, 2008). In V2Karst, the transpiration is attributed to the two first soil layers proportional to their storage content. This simple representation assumes that the roots can equally access the moisture stored in the first and second layer. Actual transpiration from the first layer $T^1_{act,i}(t)$ [mm] and the second layer $T^2_{act,i}(t)$ [mm] are computed as follows:

$$\begin{aligned}
T^1_{act,i}(t) &= \frac{V^1_{soi,i}(t-1) + T_f(t) - Es_{act,i}(t)}{V^1_{soi,i}(t-1) + T_f(t) - Es_{act,i}(t) + V^2_{soi,i}(t-1)} T^{12}_{act,i}(t),\\
T^2_{act,i}(t) &= \frac{V^2_{soi,i}(t-1)}{V^1_{soi,i}(t-1) + T_f(t) - Es_{act,i}(t) + V^2_{soi,i}(t-1)} T^{12}_{act,i}(t).
\end{aligned}$$

(~~10~~12)

~~In V2Karst, it is assumed that transpiration occurs when the canopy is dry only, as it is typically done in the other large-scale models.~~ The fraction of the day with wet canopy $t_{wet}(t)$ $[-]$ is estimated as the fraction of available energy that was used to evaporate water from the interception store ~~by assuming that the actual rate of evaporation from interception is constant throughout the day and is equal to the potential rate~~ similar to Kergoat (1998):

$$t_{wet}(t) = \frac{Ec_{act}(t)}{f_c(t) Ec_{pot}(t)}$$

(~~11~~13)

Potential evapotranspiration

We replace the Priestley-Taylor potential evaporation equation originally used in VarKarst ~~used in the previous version of the model~~ by the Penman-Monteith equation (Monteith, 1965), that has been shown to be applicable at both daily and sub-daily time step (e.g. Allen et al., 2006; Pereira et al., 2015). Potential transpiration rate over the vegetated fraction of the cell $T_{pot}(t)$ [mm] is estimated from the canopy aerodynamic resistance $r_{a,can}(\mathrm{t})$ [s m^{-1}] and surface resistance $r_{s,can}(t)$ [s m^{-1}], potential evaporation from interception over the vegetated fraction of the cell $Ec_{pot}(t)$ [mm] is assessed assuming that the surface resistance is equal to 0 following e.g. Shuttleworth (1993), while potential bare soil evaporation rate over the bare soil fraction of the cell $Es_{pot}(t)$ [mm] is calculated from the soil aerodynamic resistance $r_{a,soi}(t)$ [s m^{-1}] and surface resistance $r_{s,soi}$ [s m^{-1}], using the following equations:

$$T_{pot}(t) = \frac{\Delta(t)(R_n(t) - G(t)) + K_t\rho_a(t)c_p\dfrac{e_s(t)-e_a(t)}{r_{a,can}(t)}}{\lambda(t)\left(\Delta(t)+\gamma(t)\left(1+\dfrac{r_{s,can}(t)}{r_{a,can}(t)}\right)\right)},$$

$$Ec_{pot}(t) = \frac{\Delta(t)(R_n(t) - G(t)) + K_t\rho_a(t)c_p\dfrac{e_s(t)-e_a(t)}{r_{a,can}(t)}}{\lambda(t)\big(\Delta(t)+\gamma(t)\big)}, \qquad (\cancel{12}14)$$

$$Es_{pot}(t) = \frac{\Delta(t)(R_n(t) - G(t)) + K_t\rho_a(t)c_p\dfrac{e_s(t)-e_a(t)}{r_{a,soi}(t)}}{\lambda(t)\left(\Delta(t)+\gamma(t)\left(1+\dfrac{r_{s,soi}}{r_{a,soi}(t)}\right)\right)}.$$

w~~W~~here $R_n(t)$ [MJ m^{-2} Δt^{-1}] is the net radiation (Δt is the simulation time step), $G(t)$ [MJ m^{-2} Δt^{-1}] is the ground heat flux, $\lambda(t)$ [MJ kg^{-1}] is the latent heat of vaporization of water, Δ(t) [kPa °C^{-1}] is the gradient of the saturated vapour pressure-temperature function, $\gamma(t)$ [kPa °C^{-1}] is the psychrometric constant, $\rho_a(t)$ [kg m^{-3}] is the air density, c_p [MJ kg^{-1} °C^{-1}] is the specific heat of the air and is equal to ~~–~~ $1.013.10^{-3}$ MJ kg^{-1} °C^{-1}, $e_s(t)$ [kPa] is the saturation vapor pressure, $e_a(t)$ [kPa] is the actual vapor pressure, and K_t [s Δt~~d~~$^{-1}$] is a time conversion factor which corresponds to the number of seconds per simulation time step (equal to 86,400 s. d^{-1} for daily simulations). In this study, we neglect ground heat flux for daily time step, which seems to be a reasonable ~~for daily calculations~~assumption (see e.g. Allen et al., 1998; Shuttleworth, 2012).

The aerodynamic resistances of canopy ($r_{a,can}$(t)) and of the soil ($r_{a,soi}(t)$), that depend on the properties of the land cover and the soil respectively, are computed using the formulation of Allen (1998). To assess $r_{a,can}$(t)), roughness lengths and zero displacement plane for the canopy are estimated from the vegetation height h_{veg} [m] (Allen et al., 1998). To calculate

$r_{a,soi}(t)$), the zero plane displacement height is equal to zero ($d = 0$) and the roughness length for momentum and for heat and water vapor transfer are assumed to be equal, as in Šimůnek et al. (2009) and denoted as z_0 [m].

Finally, the canopy surface resistance is computed by scaling the stomatal resistance r_{st} [s ~~-~~m^{-1}] to canopy level using the leaf area index over the vegetated fraction (as in Eq. (5~~6~~) to assess canopy interception capacity), and therefore assuming a homogeneous response across all stomata in the canopy (Allen et al., 1998; Liang et al., 1994):

$$r_{s,can}(t) = \frac{r_{st}}{\left(\frac{LAI(t)}{f_c(t)}\right)}. \qquad (\cancel{13}15)$$

In other large-scale models, $r_{s,can}$ is also often expressed as a function of LAI, which allows to directly represents its seasonality following the variations in LAI.

Seasonality of vegetation

We represent the seasonality of vegetation by describing the seasonal variation of the cell average leaf area index LAI. We use two parameters, the maximum LAI_{max} [m^2~~-~~ m^{-2}], which is the annual maximum value of LAI during the growing season (assumed to be from June to August in this study) and LAI_{min} [%], which is the percentage of reduction in LAI during the dormant season (assumed to be from December to February in this study). The monthly value of leaf area index LAI_m [m^2~~-~~ m^{-2}] for the m^{th} month is computed using a continuous, piecewise linear function of LAI_{max} and LAI_{min}, which allows for a smooth transition between dormant and growing seasons and is similar to the function proposed by Allen et al. (1998) to assess the seasonality in crop factors:

$$\begin{aligned}
LAI_m &= \frac{LAI_{min}}{100} LAI_{max} && \text{when } m = 1, 2, 12 \\
LAI_m &= \frac{LAI_{min}}{100} \frac{LAI_{max}}{4} (6 - m) + \frac{LAI_{max}}{4} (m - 2) && \text{when } m = 3, 4, 5 \\
LAI_m &= LAI_{max} && \text{when } m = 6, 7, 8 \\
LAI_m &= \frac{LAI_{min}}{100} \frac{LAI_{max}}{4} (m - 8) + \frac{LAI_{max}}{4} (12 - m) && \text{when } m = 9, 10, 11.
\end{aligned} \qquad (\cancel{14}16)$$

The advantage of using this simple parameterisation is that it permits to easily analyse the effect of vegetation seasonality by studying the sensitivity of the model predictions to parameter LAI_{min}, which captures the strength of the seasonal variation in LAI. Timings of the four phases of the seasonality model reported in Eq. (16) are adapted for the application at the sites used in the present study, which are all located in Europe (Sect. 3.3.1).

2.3.4 Water storage in the epikarst

Epikarst water storage $V_{epi,i}(t)$ [mm] for the ith compartment is updated at the end of each time step t as follows:

$$V_{epi,i}(t) = V_{epi,i}(t-1) + R_{epi,i}(t) - Q_{epi,i}(t) - Q_{lat,\mathrm{i}\rightarrow \mathrm{i+1}}(t) \qquad \text{when } i < n_c,$$
$$V_{epi,n_c}(t) = V_{epi,n_c}(t-1) + R_{epi,n_c}(t) - Q_{epi,n_c}(t) - Q_{surf,n_c}(t) \quad \text{when } \; i = n_c.$$
(~~15~~17)

where $Q_{epi,i}(t)$ [mm] is the potential recharge to the groundwater (Eq. (18~~6~~)), $Q_{lat,\mathrm{i}\rightarrow i+1}(t)$ [mm] is the lateral flow from the ith to the $(i+1)$th model compartment and $Q_{surf,n_c}(t)$ [mm] is the surface runoff generated by the n_cth compartment. When soil and epikarst layers are saturated, the concentration flow component of the model is activated. The ith model compartment generates lateral flow towards the $(i+1)$th compartment $Q_{lat,\mathrm{i}\rightarrow i+1}(t)$ $[mm]$ equal to its saturation excess. Lateral flow from the n_cth compartment is lost from the cell as surface runoff while the other model compartments do not produce any surface runoff. The epikarst is simulated as a linear reservoir (Rimmer and Hartmann, 2012) with outflow coefficient $K_{E,i}$ [d]:

$$Q_{epi,i}(t) = \min\left(\frac{V_{epi,i}(t-1)}{K_{E,i}}, V_{epi,i}(t-1) + R_{epi,i}(t)\right).$$
(~~16~~18)

3. Sites and data for model testing

3.1 ~~Site d~~Description of study sites

We test the model with plot scale measurements from sites of the FLUXNET network (Baldocchi et al., 2001). We identified four FLUXNET sites across European and Mediterranean carbonate rock areas for which sufficient data are available to force V2Karst and to test the model (see Sect. 3.2). A short summary of each site characteristics is provided in ~~Figure~~ Fig. 2, and more detailed information can be found in Table B1.

The sites have different climate and land cover properties. The first site (Hainich site, referred to as ‘German site’) is located in the protected Hainich National Park, Thuringia, central Germany, and is characterised by a suboceanic-submountain climate and a tall and dense deciduous broadleaf forest. The second site (Llano de los Juanes site referred to as ‘Spanish site’) is located on a plateau of the Sierra de Gádor mountains, south-eastern Spain, has a semi-arid mountain Mediterranean climate and is an open shrubland. The third site (Font-Blanche site, referred to as ‘French 1 site’) is located in south-eastern France, has a Mediterranean climate and its land cover is medium-height mixed evergreen forest. The fourth site (Puéchabon site, referred to as ‘French 2 site’) is located in southern France and is characterised by a Mediterranean climate with a short evergreen broadleaf forest. Overground vegetation properties are well characterised at all sites, but subsurface properties are more uncertain. In particular, the rooting depth water capacity was only well investigated at the French 2 site. The four sites are appropriate for testing V2Karst since they satisfy the model assumptions, namely a karstified or fissured and fractured bedrock, overall high infiltration capacity with limited surface runoff and high clay content in the soil (Table B1).

3.2 Data description and preprocessing~~paration~~

Data available at the four FLUXNET sites include measurements of precipitation, temperature, net radiation, relative humidity and wind speed to force the model~~, and~~. E~~e~~ddy-covariance measurements of latent heat flux (density) and at the German and Spanish sites measurements of soil moisture are also available to estimate the model parameters (Sect. 4.1). Specifically, at the German site, soil moisture was measured in one vertical soil profile at three different depths (5, 15 and 30 cm) with Theta-probes (Knohl et al., 2003). We selected the measurement at 30 cm depth, which we deem to be most representative of the entire soil column which has a depth between 50 and 60 cm. At the Spanish site, soil moisture was assessed at a depth of 15 cm using a water content reflectometer (Pérez-Priego et al., 2013).

~~Regarding the data processing, d~~Data to force the model were gap-filled and aggregated from 30 min to daily time scale. V2Karst output observations, namely latent heat flux and soil moisture measurements, were aggregated from 30 min to monthly time scale and we discarded the months when more than 20 % of 30 min data were missing. We discarded the monthly observations of latent heat flux and soil moisture for months in which the forcing data contain many gaps, and therefore the impact of the gap-filling of the data on the simulation results is likely to be too significant to sensibly compare simulated and observed soil moisture and latent heat flux. Additionally, we removed monthly aggregated latent heat flux measurements when the mismatch in the energy balance closure was higher than 50% similar to (Miralles et al., 2011). We corrected latent heat flux measurements to force the closure of the energy balance assuming that measured latent heat flux LE $[\mathrm{MJ.\,m^{-2}.\,month^{-1}}]$ and sensible heat flux H $[\mathrm{MJ\,m^{-2}\,month^{-1}}]$ have similar errors. The corrected evapotranspiration estimate $E_{act,cor}$ $[\mathrm{mm.\,month^{-1}}]$ is calculated as (Foken et al., 2012; Twine et al., 2000):

$$E_{act,cor} = \frac{R_n}{\lambda(1+\frac{H}{LE})} \tag{19}$$

Further details on the data processing and on the analysis of the uncertainty in latent heat flux measurements are ~~is~~reported in Sect. S4 of our Supplementary material. In particular, we showed that, while the actual value of latent heat flux can have large uncertainties, we have a much higher confidence in its temporal variations.

Table 2 reports the simulation period and the number of monthly latent heat flux and soil moisture observations that were used to estimate the model parameters at the four FLUXNET sites. We extracted a continuous time series of forcing data covering about 10 years at the German site, 7 years at the Spanish site, 3 years at the French 1 site and 8 years at the French 2 site, while latent heat flux and soil moisture measurements ~~are~~ were not available over the entire simulation time series. All model simulations were performed using a one-year warm-up period, which we found to be sufficient to remove the impact of the initial conditions on the simulation results (see Sect. S5 of our Supplementary material).

~~Moreover, we corrected latent heat measurements and analysed their uncertainty. We derived two corrected estimates of actual ET, obtained by forcing the closure in the energy balance following (Foken et al., 2012; Twine et al., 2000), namely:~~

1. ~~a corrected value that assumes that latent heat (LE [$MJ.m^{-2}.month^{-1}$]) and sensible heat (H [$MJ.m^{-2}.month^{-1}$]) have similar errors (referred to as Bowen ratio estimate):~~

$$\cancel{E_{act,bow} = \frac{R_n}{\lambda.(1+\frac{H}{LE})} [mm.month^{-1}],} \quad (17)$$

2. ~~a corrected value that assumes errors in latent heat only (referred to as residual estimate):~~

$$\cancel{E_{act,res} = \lambda.(R_n - H) [mm.month^{-1}].} \quad (18)$$

~~An additional analysis showed that the two corrected estimates of Eq. (17-18) and the uncorrected measure of actual ET are well correlated at the FLUXNET sites, which gives us some confidence regarding the temporal variations in actual ET measurements, while relative errors between corrected and uncorrected estimates can be quite high (see Section S4 of our Supplementary material). We chose to use the Bowen ratio estimate (Eq. (17)) to calibrate the model. In fact, it is not clear whether one of the two turbulent fluxes may be more uncertain than the other (Foken et al., 2012).~~

4. Methods

~~In this study, we estimate V2Karst parameters and test the plausibility of model realisations at the FLUXNET sites (Sect. 4.1), we conduct a global sensitivity analysis of the model parameters at the FLUXNET sites to identify the model dominant controls and inform model calibration for future applications (Sect. 4.2) and we last, perform a set of virtual experiments to learn about the mechanism of recharge production in the model and its sensitivity to precipitation characteristics and land cover type (Sect. 4.3).~~ To test the plausibility of V2Karst predictions at FLUXNET sites, we run three sets of analyses: we first estimate the model parameters using actual ET and soil moisture observations available at the four sites (methods in Sect. 4.1). We then analyse the sensitivity of simulated annual recharge to the model parameters using measured forcing data to understand the controlling processes (Sect. 4.2). Finally, we investigate the sensitivity of simulated recharge to precipitation properties and land cover type in a virtual experiment to understand the controlling processes in recharge when the forcing data is varied more widely than observed at the study sites (Sect. 4.3). All the analyses were performed using the SAFE toolbox for global sensitivity analysis (Pianosi et al., 2015).

4.1 Parameter ~~estimation~~ constraining at the FLUXNET sites using soft rules

We investigate whether it is possible to estimate parameter values that produce plausible simulations based on information available at each FLUXNET site. To this end, and similarly to Hartmann et al. (2015), we use 'soft rules' to accept or reject parameter combinations based on the consistency between monthly model simulations on one side, and monthly observations and a priori information on model fluxes on the other side. Using soft rules instead of 'hard rules' (i.e. minimisation of the mismatch between observations and simulations) allows to identify a set of plausible ~~model simulations~~ parameter sets and accounts for the fact that (1) the observed soil moisture is not strictly commensurate with simulated soil moisture, (2) observations are affected by uncertainties (see Sect. 3.2) and (3) it is not expected that V2Karst simulations closely match site-specific data, since the model structure is based on general understanding of karst systems for large-scale applications and may not account for some site specificities. We define five soft rules to identify acceptable ('behavioural') parameter combinations:

1. **The bias between observed and simulated actual ET is below 20**%:

$$Bias = \left|\frac{\sum_{t\in M_{ET}}(E_{act,sim}(t) - E_{act,\cancel{bow}cor}(t))}{\sum_{t\in M_{ET}} E_{act,\cancel{bow}cor}(t)}\right| < 20\%, \quad (~~19~~20)$$

   where $E_{act,sim}(t)$[mm] is the simulated actual ET for month t (sum of transpiration, soil evaporation and evaporation from canopy interception), $E_{act,\cancel{bow}cor}(t)$ [mm] is the ~~Bowen ratio correction of~~corrected observed actual ET (Eq. (~~17~~19)), and M_{ET} is the set of months for which latent heat measurements are available. This rule allows to constrain the simulated water balance.

2. **The correlation coefficient** (ρ_{ET}) **between observed monthly actual ET** *($E_{act,bw}$)* **and simulated total actual ET** ($E_{act,sim}$) **is above 0.6.** This rule ensures that the temporal pattern of simulated ET follows the observed pattern.

3. **The correlation coefficient** (ρ_{SM}) **between observed monthly soil moisture** (SM_{obs} [% soil saturation]) **and simulated monthly soil moisture** (SM_{sim} [m^3 ~~.~~m^{-3} soil volume]) **is above 0.6.** Simulated soil moisture SM_{sim} for month t is calculated as the average soil moisture within the root zone over all model compartments. This rule guarantees that soil moisture variations are consistent with observations.

4. **Total simulated surface runoff** (Q_{surf}) **is less than 10% of precipitation**, in accordance with a priori information on the carbonate rock sites, which attests that runoff is negligible (see section 3.1).

5. **Soil and vegetation parameter values are consistent with a priori information**, i.e. they fall within constrained (site-specific) ranges. This rule applies to the parameters for which a priori information is available at the FLUXNET sites, namely h_{veg}, r_{st}, LAI_{min}, LAI_{max} V_r and V_{soi} and the constrained ranges are reported in Table 3. This rule ensures that acceptable model outputs are produced using plausible parameter values.

For each site, we derived a sample of size 100,000 for the 15 parameters of V2Karst using Latin hypercube sampling and unconstrained (wide) ranges for the model parameters to explore a large range of soil and vegetation types. We applied the above rules in sequence to either reject or accept the sampled parameter combinations. We sampled more densely the constrained parameter ranges used in rule 5 so that a sufficiently large number of parameterisations remain after applying rule 5. Similarly to Hartmann et al. (2015), a priori information on parameter ranges (rule 5) is applied last so that we can first assess the constraining of the parameter space based on information on model output only (rules 1 to 4), and then the consistency of this constraining with a priori information (rule 5).

We also note that the thresholds used in rules 1 to 3 are stricter compared to the study by Hartmann et al. (2015), in which the threshold for the bias rule (1) was set to 75% and for the correlation rules (2 and 3) was set to 0. The reason is that in

Hartmann et al. (2015) behavioural parameter sets had to be consistent with observations at all sites within each climate zone defined in the study, while here we perform the parameter estimation for each site separately and therefore we expect better model performances.

4.2 ~~Parameter g~~Global sensitivity analysis ~~of the model parameters~~

We use a global sensitivity analysis called the Elementary Effect Test (Saltelli et al., 2008), also sometimes referred to as~~or~~ method of Morris (Morris, 1991), ~~. This is a global sensitivity analysis method, and therefore it permits~~ to analyse sensitivity across the entire space of parameter variability ~~space~~ (Saltelli et al., 2008). We chose this method in particular because it is well suited for identifying uninfluential parameters (Campolongo et al., 2007; Saltelli et al., 2008), which is one of the objectives of our analysis, and because it can be ~~easily~~ applied to dependent parameters (in V2Karst it is assumed that $V_e \leq V_r \leq V_{S,n}$ as explained in Sect. 2.3.1).

The method requires the computation of the Elementary Effects (EEs) of each parameter in n different baseline points in the parameter space. The EE of the ith parameter x_i at given baseline point $\left(x_1^j, x_2^j, \ldots, x_{i-1}^j, x_i^j, \ldots, x_M^j\right)$ and for a predefined perturbation Δ is assessed as follows:

$$EE_i^j = \frac{y\left(x_1^j, x_2^j, \ldots, x_{i-1}^j, x_i^j + \Delta, \ldots x_M^j\right) - y\left(x_1^j, x_2^j, \ldots, x_{i-1}^j, x_i^j, \ldots x_M^j\right)}{\Delta},$$

(2~~0~~1)

where M is the number of parameters and y is the model output (simulated annual recharge in our case). ~~The sensitivity indices analysed in the present study are~~ We analyse the mean of the absolute values of the EEs (denoted by μ_i^*) introduced in Campolongo et al. (2007), which is a measure of the total effect of the ith parameter, and the standard deviation of the EEs (σ_i) proposed in (Morris, 1991), which is an aggregate measure of the intensity of the interactions of the ith parameter with the other parameters and of the degree of non-linearity in the model response to changes in the ith parameter.

The total number of model evaluations required to compute these two sensitivity indices is $n(M+1)$, where n is the number of baseline points chosen by the user. The baseline points and the perturbation Δ of Eq. (~~20~~21) were determined following the radial design proposed by Campolongo et al. (2011). The baseline points were randomly selected using Latin hypercube sampling for the 15 parameters of V2Karst, and dropping the parameter sets that did not meet the condition $V_e \leq V_r \leq V_{S,n}$. In our application, we used $n = 500$ points, which means that we needed 8000 model evaluations for each sensitivity analysis for each of the four FLUXNET sites. We derived confidence intervals on the sensitivity indices ~~via bootstrapping~~ using 1000 bootstrap resamples, and checked the convergence of the results at the chosen sample size, as suggested in Sarrazin et al. (2016).

4.3 Virtual experiments to analyse sensitivity to climate and land cover change

Our last analysis consists of a set virtual experiments to investigate the sensitivity of recharge~~ and actual ET ~~simulated by V2Karst to changes in (1) the precipitation properties, ~~(~~specifically monthly total precipitation ~~average monhly amount~~ and daily temporal distribution, i.e. frequency and intensity, which are likely to change (IPCC, 2013)~~)~~ and (2) land cover (specifically from forest to shrub and vice versa).

Virtual experiments ~~using numerical models~~permit full control on experimental conditions, and thus to unequivocally attribute changes in model outputs to changes in model inputs (see e.g. Pechlivanidis et al., 2016; Weiler and McDonnell, 2004). ~~Several~~ Different studies have used virtual experiments to analyse the impact of precipitation spatial and temporal variability on hydrologic model outputs. In fact, using historical precipitation time series or future projections only allow exploration of a limited range of possible realisations, which makes it difficult to disentangle the effects of different precipitation properties on model outputs. Instead, synthetic precipitation time series can be tailored to analyse the impact of specific precipitation characteristics, for instance precipitation spatial distribution (Pechlivanidis et al., 2016; Van Werkhoven et al., 2008) and precipitation temporal distribution, namely frequency and intensity (Jothityangkoon and Sivapalan, 2009; Porporato et al., 2004), storminess (Jothityangkoon and Sivapalan, 2009) and seasonality (Botter et al., 2009; Jothityangkoon and Sivapalan, 2009; Laio et al., 2002; Yin et al., 2014).

In this study, we create a synthetic precipitation time series where the same precipitation event is periodically repeated. The precipitation time series is characterized by the intensity of precipitation events I_p $[\mathrm{mm}\,\mathrm{d}^{-1}]$ and the interval between two wet days H_p $[\mathrm{d}]$. The duration of each precipitation event here is set to one day. The average monthly precipitation P_m $[\mathrm{mm.\,month}^{-1}]$ for an average month with 30 days is therefore equal to:

$$P_m = 30.\frac{I_p}{1 + H_p}$$ (~~21~~22)

To determine the possible range of variation of the three variables, P_m, I_p and H_p, we analysed their distributions at the four FLUXNET sites and over all European and Mediterranean carbonate rock areas using GLDAS data (Rodell et al., 2004) (distributions are reported in section S6 of our supplementary material). We found that wide but plausible ranges are: P_m varies between 0 and 500 $\mathrm{mm\,month}^{-1}$, I_p varies between 0 and 200 $\mathrm{m\,d}^{-1}$ and H_p varies between 0 and 89 d (note that $H_p = 0$ means that it rains every day). We then derived a set of 2266 precipitation time series by deterministically sampling P_m, and H_p within those ranges (and consequently deriving a sampled value of I_p from Eq. (~~21~~22)). We sampled more densely closer to the lower bound of the ranges since lower values of P_m and H_p are more likely to occur. For each of the precipitation time series so obtained, we ran the V2Karst model until ~~the simulated fluxes reached~~ a steady-state is reached (i.e. periodic oscillations of all state and flux variables) and we analyse the steady-state monthly average of recharge~~, transpiration, soil evaporation and evaporation from interception.~~.

The experiments are conducted at two 'virtual sites' that are designed based on the characteristics of the FLUXNET sites. Specifically, we use a virtual 'forest site' that has the characteristics of the German site as the baseline (i.e. its behavioural parameterisations for the soil, epikarst and vegetation parameters, and its climate characteristics) and a virtual 'shrub site' that has the characteristics of the Spanish site. We do not investigate the effects of varying temperature, net radiation, relative humidity and wind speed characteristics as we did for precipitation, because these variables are correlated (see e.g. Ivanov et al., 2007) and therefore they cannot be varied independently. We account for their overall combined effect and we vary them jointly to reproduce two conditions, i.e. winter (e.g. low energy for ET) and summer (high energy for ET). For winter (respectively summer) conditions we forced the model using constant values of temperature, net radiation, relative humidity and wind speed that were taken as the average values measured at the FLUXNET sites during winter (respectively summer). ~~The forest site also inherits the suboceanic-submountain climate characteristics of the German site (i.e. we force the model by the average values of air temperature, net radiation, humidity and wind speed measured at that site), while the shrub site inherits the semi-arid climate of the Spanish site.~~ To investigate the impact of a change in land cover ~~at these virtual sites~~, we swapped the vegetation parameters ~~(indicated in Table 1)~~ between the two virtual sites.

~~We do not investigate the effects of varying temperature, net radiation, relative humidity and wind speed characteristics as we did for precipitation, because these weather variables are correlated (see e.g. (see e.g. Ivanov et al., 2007)) and therefore they cannot be varied independently. Instead, we account for their overall combined effect in a simple way by analysing the changes in sensitivity when these variables are set to winter (low energy for ET) and summer (high energy for ET) conditions.~~ Table 4 reports the values of the parameters and the values of the weather variables used at the two virtual sites.

5. Results

5.1 Parameter ~~estimation~~constraining at FLUXNET sites

~~In this section, we present the results of the parameter estimation at FLUXNET sites. We analyse the impact of the application of the soft rules defined in Sect. 4.1 on the reduction in acceptable ('behavioural') parameterisations (Sect. 5.1.1) and we examine V2Karst outputs (Sect. 5.1.2)~~

~~5.1.1 Analysis of the constraining of the parameter space~~

We first present the results of applying the soft rules defined in Sec. 4.1 at the four FLUXNET sites. Figure 3 shows that behavioural parameterisations consistent with all rules can be identified at all sites~~, but~~ T~~t~~heir number is very different from one site to another. Specifically, out of the initial 100,000 randomly generated parameter samples, we found 36,838 behavioural parameterisations at the German site, 147 at the Spanish site, 6354 at the French 1 site and 4077 at the French 2 site. From Fig. 3, we also see that the application of each rule reduces the number of behavioural parameterisations, except for rule 4 (value of total surface runoff < 10% of precipitation), since all model simulations produce less than 7% of surface runoff at all sites.

This can be explained by the fact that V2Karst gives priority to recharge production over surface runoff. Therefore, the latter only occurs under extremely wet conditions when all model compartments are saturated.

Figure 4 reports a parallel coordinate plot of the behavioural parameter sets and associated values of the output metrics after sequential application of the soft rules. The application of rules 1 to 4 does not significantly reduce the parameter ranges, but it only allows to discard low values of parameters V_r and V_{soi} at all sites (dark blue lines in Fig. 4). Instead, the application of rule 5 (a priori parameter ranges, ~~red~~ black vertical lines in Fig. 4) permits a significant reduction in parameter ranges, not only for the parameters that are directly constrained by this rule (h_{veg}, r_{st}, LAI_{min}, LAI_{max}, V_r and V_{soi}) but also for the spatial variability coefficient a. Specifically, behavioural values of parameter a are found to be between 0 and 3.2 at the French 1 site, between 0 and 2.8 at the French 2 site. At the Spanish site, we also observe that the behavioural simulations (red lines) cover more densely some portions of the ranges, specifically higher values of parameters $r_{s,soi}$ and a, and lower values of z_0 and V_e. This means that the value for these parameters is more likely to be within these sub-ranges.

~~5.1.2 Analysis of model simulations~~predictions

~~In this paragraph,~~ W~~w~~e also analyse the repartition of the simulated water fluxes when using the behavioural parameterisations ~~(Fig. 5)~~. The top panel (~~Figure~~ Fig. 5a) compares the total simulated recharge and the total actual ET, expressed in percentage of total precipitation at the four FLUXNET sites (mean and 95% confidence interval across the behavioural parameterisations). At the Spanish site, we present the results over two different time periods that have very different precipitation amounts, namely a drier period from 1 January 2006 to 31 December 2008 and a wetter period from 1 January 2009 to 30 December 2011 (see Fig. 2). Figure 5 shows that, apart from extremely wet periods at the Spanish site, in all other cases the fraction of recharge (Q_{epi}) is significantly lower than the fraction actual ET (ET_{act}). Figure 5b shows the partitioning of ET among its different components (transpiration, soil evaporation and interception). We observe that transpiration (T_{act}) is the largest component at all sites, while the relative importance of evaporation from canopy interception (Ec_{act}) and soil evaporation (Es_{act}) varies across sites. In particular, at the German site, Ec_{act} is on average particularly high compared to the other sites~~, which may be partly explained by the fact that summer LAI (parameter LAI_{max}) is higher at this densely forested site compared to the other sites, and therefore the summer canopy storage capacity is higher as well.~~.

Finally, Fig. 6 ~~presents~~ reports monthly time series of both observed variables, namely precipitation input P, actual $E_{act,cor}$ (in blue) and soil moisture SM_{obs} (in green). It also reports simulated variables, namely recharge Q_{epi}, actual ET $E_{act,sim}$ and soil moisture in the root zone SM_{sim} (non-behavioural simulations are in grey and behavioural simulations in black). ~~the time series of monthly precipitation input (P), simulated monthly recharge (Q_{epi}), total actual ET (E_{act}) and soil moisture in the root zone (SM_{sim})~~ at the four FLUXNET sites. ~~Observation of soil moisture and actual ET are also reported, and the blue lines correspond to the Bowen ratio corrected estimate used in rules 1-2 for parameter estimation (see Sect. 4.1)~~. We see that the soft rules allow to significantly reduce the uncertainty in model outputs at all sites. In fact, the width of the behavioural ensemble, i.e. the ensemble of simulations obtained by application of the rules (black lines), is much narrower than the non-behavioural ensemble (grey lines). Simulated actual ET ($E_{act,sim}$) is also closer to the observations (blue line) in the behavioural ensemble compared to the non-behavioural one. This means that the application of the soft rules and a priori information on parameter ranges allows not only to improve the precision of the simulated states and fluxes (reduced uncertainty ranges of the simulations), but also the accuracy of simulated actual ET (simulations close to observations). ~~Moreover, the model structure is flexible enough to capture most corrected and uncorrected ET observations, since the non-behavioural model ensemble (grey) includes most corrected and uncorrected ET values.~~

From Fig. 6, we also observe that the seasonal variations in model predictions are consistent with our understanding of the sites over the entire simulation horizon and not only over the months for which ET and soil moisture observations are used to estimate the parameters (blue and green areas in the plot). Specifically, at the German site we find a marked seasonality of simulated $E_{act,sim}$ and SM_{sim}, with low $E_{act,sim}$ and high SM_{sim} in winter, and high $E_{act,sim}$ and low SM_{sim} in spring and summer. In fact, in winter, the energy available for ET is low and the deciduous vegetation is not able to transpire or intercept large amounts of precipitation, while in spring and summer more energy is available for ET and the vegetation has a higher value of LAI, and therefore ET losses can occur and deplete the soil moisture. At the other sites we observe a similar pattern for SM_{sim}, while E_{act} tends to peak in spring and to be lower in summer when the ET fluxes are more water-limited than at the German site.

5.2 ~~Parameter g~~Global sensitivity analysis of the model parameters

The sensitivity analysis results are reported in Fig. 7 and refer to the sensitivity of total simulated recharge (expressed as a percentage of total precipitation) to the 15 parameters of the V2Karst model. For each parameter, the plots in Fig. 7 report on the horizontal axis the absolute mean of the Elementary Effects (μ^*, total effect of the parameters) and on the vertical axis their standard deviation (σ, degree of linearity of the effect of the parameters). In all plots, we observe that the bootstrap confidence intervals of the sensitivity indices are narrow and show little overlap, which gives confidence that the sensitivity results are robust. Similar to the analysis of the simulated fluxes in Sect. 5.1.2 (Fig. 5), at the Spanish site we present the results for two different time periods with different precipitation amounts. ~~.~~

~~5.2.1 Global sensitivity analysis with constrained parameter ranges~~

We first examine the left panels in Fig. 7, which show the sensitivity results when ~~(~~h_{veg}, r_{st}, LAI_{min}, LAI_{max}, ~~and~~ V_r~~) and the soil storage capacity~~ V_{soi} are sampled within constrained ranges (Table 3), while the ranges for other parameters are taken from Table 1. The objective is to inform model calibration in future model applications, since such parameter ranges capture the uncertainty in parameter values left after considering site-specific information. We first note that μ^* and σ take a non-zero value for all parameters at all sites, which means that all parameters are influential and have a non-linear effect on recharge, possibly through interactions with other parameters. ~~The existence of parameter interactions can explain the limited reduction in some parameter ranges during our parameter estimation (Sect. 4.1).~~

We observe that the spatial variability coefficient a has by far the largest influence, followed by parameters V_{soi} and V_r. In fact, their value of μ^* is significantly higher than the other parameters at all sites. The implication for model calibration in future applications of V2Karst is that efforts should primarily seek to reduce the uncertainty in parameters a, V_{soi} and V_r. These three parameters also have a significantly large value of σ, which indicates non-linearities in the model response to variations in these parameters~~ and which is coherent with the nature of Eq. (1-2)~~. Interestingly, parameter V_{can}, that controls evaporation from interception, and $r_{s,soi}$, that controls soil evaporation, have an impact on recharge at most sites and at the Spanish site during wet years respectively. This shows that the processes of evaporation from interception and soil evaporation can be important for recharge simulations.

Moreover, we observe that parameters LAI_{min}, z_0 k and V_e have a very small impact on total recharge at all sites ($\mu^* < 3$ %). However, Section S7 of our Supplementary material reports additional sensitivity analysis results for other model outputs and shows that the most influential parameters that should be the focus of the calibration strategy vary depending on the output of interest. ~~In particular, parameter V_e has a significant impact on the fraction of actual transpiration in total ET, and therefore on the partitioning of ET among its different components.~~

~~**5.2.2 Global sensitivity analysis with unconstrained parameter ranges**~~

We then examine ~~The~~ the right panels of Fig. 7, which shows the sensitivity indices when sampling parameters within unconstrained ranges (Table 1). This analysis allows to test the plausibility of the model structure through the assessment of the model sensitivity across a large spectrum of soil and vegetation conditions.

The most apparent difference with respect to the previous results (Sect. 5.2.1) is that vegetation parameters (h_{veg}, r_{st}, LAI_{min} LAI_{max} and V_r) now have a much higher value of the sensitivity indices (both μ^* and σ). More specifically, LAI_{max} has a very high sensitivity index at all sites ($\mu^* > 10.5\%$), which can be attributed to the fact that this parameter is used to calculate different model components. Interestingly, the seasonality of leaf area index appears to play an important role in V2Karst since μ^* for LAI_{min} , although always lower than μ^* for LAI_{max} , stands out at all sites.

~~When comparing~~ The considerable variations in parameter sensitivities observed across sites~~, we see some significant differences, that we~~ can be interpreted by considering their climatic differences. ~~In fact, we~~We would expect transpiration to be mainly energy-limited at the German site, given that it has a suboceanic-submountain climate and mainly water-limited at the French sites, which have a Mediterranean climate, and at the Spanish site, which has a semi-arid Mediterranean climate. ~~Specifically~~In this regard, the most influential parameter at the Spanish site is by far parameter a (high μ^*), which has an impact on the water storage in the soil and therefore on the amount of water available to sustain ET between rain events, while at the German site~~, parameter LAI_{max} , which is used to calculate PET, has the largest effect on recharge (high μ^*).We also notice that~~ parameters r_{st} and h_{veg}, that control PET, are more influential ~~at the German site~~ compared to the other sites.

Finally, we observe that, the parameters that specifically control the volume of transpiration (r_{st} and V_r) have a significantly higher value of μ^* than the parameters that specifically control soil evaporation (z_0, $r_{s,soi}$ and V_e) and evaporation from interception (V_{can}). Moreover, z_0, $r_{s,soi}$ and V_e have a very small impact ($\mu^* < 3\,\%$), while parameter V_{can} can have an important effect at the German site ($\mu^* = 5.7\,\%$). ~~This suggests that transpiration is overall dominating the ET fluxes at these sites when exploring a wide range of soil and land cover properties and that interception is an important process under the climate of the German site.~~ Additionally, we see that parameter f_{red}, that controls transpiration from the third soil layer, has an impact on recharge simulated at the Spanish site.

5.3 Virtual experiments ~~to analyse sensitivity to precipitation and land cover characteristics~~

After showing that the V2Karst model behaves reasonably at the four FLUXNET sites, in this section we use virtual experiments to further learn about the sensitivity of simulated recharge to precipitation characteristics and land cover using virtual sites (see Sect. 4.3).

~~5.3.1 Sensitivity of simulated fluxes to precipitation characteristics~~

Figure 8 shows the monthly average value of simulated recharge Q_{epi}, for a range of synthetic precipitation inputs with different values of the precipitation monthly amount P_m (x-axis) and of the interval between rainy days H_p (y-axis) at the virtual forest and shrub sites. We do not report Q_{epi} values in the top right of the plots because this region corresponds to very intense precipitation events (higher than 200 mm.d^{-1}) that have a very low probability of occurrence (see Sect. 4.3).

From the top left panel of Fig. 8, we see that winter Q_{epi} is mostly sensitive to P_m, in fact simulated recharge increases when moving along the horizontal direction from left to right~~,~~ but shows little variations along the vertical direction (when H_p is varied). This ~~resuls~~ result is due to the fact that actual ET is very limited in winter because of the low energy available. We indeed estimated that the maximum value of total ET across the different precipitation inputs is 13 mm ~~.~~month^{-1} at the forest site and 35 mm month^{-1} ~~mm.month^{-1}~~ at the shrub site. Therefore, a large part of precipitation becomes recharge rather independently of its temporal distribution.

From the right panel of Fig. 8, we observe a systematic reduction in summer Q_{epi} compared to winter at both virtual sites. Moreover, summer recharge is overall highly sensitive not only to P_m but also to H_p, since it increases when moving along the vertical direction from bottom to top, i.e. when the same amount of monthly precipitation falls in less frequent but more intense events. This result can be explained by the fact that in summer potential ET is larger and therefore, if events are less intense, a larger part of the precipitation is lost via ET.~~,~~ ~~while if~~If instead events are more intense, the canopy and soil stores reach saturation. In this case, ~~and precipitation generates~~ a saturation excess flow to the epikarst is generated and hence more recharge and less ET. Moreover, in summer, Q_{epi} shows a limited sensitivity to P_m and H_p when these quantities take low values (brown and red dots on the left of the plots), because only few soil compartments reach saturation under drier conditions and therefore little recharge can be generated. We also see that at the shrub site, Q_{epi} is a significant flux ($Q_{epi} > 5mm$) for smaller values of P_m and H_p compared to the forest site, which may be due to the fact that at the shrub site, the soil water capacity (V_{soi}) is much smaller and therefore the soil compartments can reach saturation under drier conditions.

~~5.3.2 Sensitivity of simulated fluxes to land cover change~~

Finally, Fig.~~ure~~ 9 reports the results of another virtual experiment ~~similar to Fig. 8~~ but focusing on the impact of land cover change. Specifically, the panels in Fig. 9 show the variation in simulated recharge when land cover is changed from forest to shrub at the virtual forest site (and vice versa at the virtual shrub site), and more specifically, Fig. 9 reports $\Delta Q_{epi} = Q_{epi}^{shrub} - Q_{epi}^{forest}$. We see that in all plots ΔQ_{epi} is positive, which means that recharge is larger and therefore actual ET is lower under shrub compared to forest land cover for both sites. From the left panels of Fig. 9, we observe that ΔQ_{epi} is very limited in winter, which is expected since ET fluxes are small in winter as explained in Sect. 5.3.1.

Instead, the right panels of Fig. 9 show that summer ΔQ_{epi} is much higher compared to winter conditions and is sensitive to both P_m and H_p. ~~The value of summer ΔQ_{epi} is largest when the monthly precipitation P_m is high and the interval between wet days H_p is low (green dots at the virtual forest site and dark blue dots at the virtual shrub site), because under these precipitation conditions the amount of moisture available for ET is maximum. Interestingly, for both virtual sites, summer ΔQ_{epi} is sensitive to both P_m and H_p, but its sensitivity~~ The sensitivity of ΔQ_{epi} is highly variable across the different precipitation inputs, and more specifically an increase in H_p can have a different effect on ΔQ_{epi} depending on the value of P_m (no variation, increase or decrease in ΔQ_{epi}). In fact, when P_m is low, ΔQ_{epi} is always low and does not vary sensibly when P_m and H_p are varied (brown area in the left end of the plot), since recharge is always low under these precipitation conditions as shown in Fig. 8. For intermediate values of P_m, ΔQ_{epi} has a similar pattern at both sites and increases when either H_p or P_m increases. Instead, for high values of P_m, we see that for both sites ΔQ_{epi} decreases when H_p increases ~~and that at the virtual forest site, ΔQ_{epi} increases when P_m increases~~. ~~The value of summer~~ ΔQ_{epi} is largest when the monthly precipitation P_m is high and the interval between wet days H_p is low (green dots at the virtual forest site and dark blue dots at the virtual shrub site), because under these precipitation conditions the amount of moisture available for ET is maximum.

Importantly, our results also show that the impact of a change in land cover can vary greatly across sites, since at the virtual shrub site summer ΔQ_{epi} reaches much higher values and is sensitive to P_m and H_p over a larger range of values of P_m and H_p compared to the virtual forest site.

6 Discussion

6.1 Plausibility of V2Karst simulations against site specific information

We tested the model by evaluating its ability to reproduce observations at four carbonate rock FLUXNET sites, which is a standard approach to model testing, used for instance to test the previous version of the model VarKarst (Hartmann et al., 2015) and large-scale ET products (Martens et al., 2017; McCabe et al., 2016; Miralles et al., 2011). We demonstrated that V2Karst is able to produce behavioural simulations consistent with observations and a priori information at FLUXNET sites. In addition, ~~, and additionally~~ ~~that~~ the time series of ~~the model outputs~~simulated water balance components are coherent with our understanding of the sites. A different number of behavioural parameterisations was identified at the different sites, ~~because we used the same constrains across sites. The fact that~~with the highest number ~~of behavioural parameterisations was~~ found at the more humid German site and the lowest number at the semi-arid Spanish site, which is coherent with previous findings that ~~higher~~ a better fit-to-observation can be obtained at wetter locations (Atkinson et al., 2002; Bai et al., 2015).

Interestingly, the behavioural parameter~~s~~ sets for the French 1 site~~, the results of the parameter estimation~~ ~~allow to~~ corroborate the hypothesis that root water uptake is likely to extent below the physical soil layer, as communicated by Guillaume Simioni (investigator of the site). In fact, we found ~~here that~~ behavioural values of parameter V_r ~~are higher than~~above 59 mm, while site-specific information indicates that the physical soil layer has a storage capacity of 49 mm (Table B1). This result further attests to the realism of V2Karst structure.

~~Moreover,~~ The results of the global sensitivity analysis of simulated recharge to the model parameters is well interpretable in light of the different climatic conditions at the four FLUXNET sites. Parameters that control PET and the soil water storage capacity generally have a large impact on simulated recharge. The interception capacity is also very important ~~(~~at the German site~~)~~. These findings are in line with the few sensitivity analysis studies of ~~-~~large-scale hydrological models~~,~~ (Güntner et al., 2007; Werth et al., 2009), ~~-~~which examined the sensitivity of continental water storage and river discharge simulated by the WaterGap model. The importance of the maximum leaf area index and, to a lesser extent, of the seasonality of vegetation in the V2Karst model is also consistent with previous studies. For example, Tesemma et al. (2015) found that assimilating year-to-year monthly LAI in the VIC model can significantly improve runoff simulations compared to using long-term average LAI, and Rosero et al. (2010) determined that LAI has a large influence on simulated latent heat in the Noah land surface model. ~~This~~ Overall, our sensitivity analysis results therefore suggest~~s~~ that the model behaves sensibly and consistently with our understanding of the key vegetation-recharge processes we aim at reproducing.

6.2 Sensitivity of simulated groundwater recharge to changes in climate and vegetation characteristics ~~in karst areas~~

~~In this study, we investigated the sensitivity of simulated recharge to both climate and land cover change, through a global sensitivity analysis of the model parameters at the FLUXNET sites, and through virtual experiments using a simple synthetic periodic precipitation input.~~

~~Firstly, the results of Elementary Effect Test using unconstrained (wide) ranges showed that the vegetation parameters (h_{veg}, r_{st}, LAI_{min}, LAI_{max} and V_r and additionally V_{can} at the German site) have a significant impact on simulated recharge at the FLUXNET sites, which means that simulated recharge is sensitive to changes in land cover properties. More specifically, the maximum leaf area index (LAI_{max}) was highly influential at all sites, and to a lesser extent the parameter controlling the seasonality in LAI (LAI_{min}). This is consistent with the findings of previous studies, since (Tesemma et al., 2015) found that assimilating year-to-year monthly LAI in the VIC model can significantly improve runoff simulations compared to using long-term average LAI and (Rosero et al., 2010) determined that LAI has a large influence on simulated latent heat in the Noah land surface model. Therefore, the future potential increasing trend in global LAI documented by (Zhu et al., 2016) could have a significant impact on the partitioning between green and blue water, including in karst areas.~~

~~Our results are also comparable to the sensitivity analysis results obtained for the WaterGap model in (Güntner et al., 2007) and (Werth et al., 2009) with respect to continental water storage and additionally runoff for the latter study. These two studies are the only ones to the author knowledge that performed a parameter global sensitivity analysis including land cover parameters for the large-scale models of Table A1. Similar to our results, both studies found that highly influential parameters are parameters that control PET (Priestley-Taylor empirical coefficient in WaterGap, which is replaced by parameters r_{st} and h_{veg} in V2Karst), the water storage capacity in the root zone (denoted as V_r in V2Karst) and at a few sites the interception capacity per unit of LAI (denoted as V_{can} in V2Karst). We note that the impact of parameter LAI was not reported and vegetation seasonality was not considered in these two studies.~~

~~Secondly, the results of our virtual experiment showed that simulated recharge is sensitive not only to changes in the precipitation amount but also in the precipitation temporal distribution (interval between wet days) and in land cover, and that its sensitivity is highly dependent on the precipitation properties and on the value of the other weather variables that are used to calculate PET (temperature, net radiation, relative humidity and wind speed). These findings indicate that it is critical to assess the combined impact of changes in all these variables on karst groundwater recharge to gain insights on future water availability in karst areas. A previous study by (Hartmann et al., 2017) also found that recharge simulated with VarKarst is sensitive to the precipitation amount and temporal distribution (specifically intensity of heavy precipitation events), using historical weather time series. Here we complemented the study of (Hartmann et al., 2017) by unequivocally attributing the changes in recharge to changes in precipitation properties using virtual experiments. Our results are also consistent with past studies for non-karst areas that established dependencies between hydrological fluxes on one side and precipitation properties on the other side, using synthetic precipitation inputs (Jothityangkoon and Sivapalan, 2009; Porporato et al., 2004) and~~

~~observations of recharge in a semi-arid tropical region (Taylor et al., 2013) and in a seasonally humid tropical region (Owor et al., 2009), and comparing different approaches for the temporal disaggregation of projected monthly precipitation to daily values to force the WaterGap model (Weiß and Alcamo, 2011). However, to the author knowledge, no previous study had systematically examined the combined impact of changes in specific precipitation characteristics and in land cover on the water balance.~~

Our results also provide valuable insights about possible vulnerabilities of karst systems. For example, the fact that LAI is a highly influential parameter at all FLUXNET sites suggests that the increasing trend in global LAI documented by Zhu et al. (2016) could reflect into a significant impact on groundwater recharge and the partitioning between green and blue water in karst areas.

Variations in simulated recharge were also systematically examined under combined changing land cover type and climate conditions (precipitation overall amount and temporal distribution), since the model should be appropriate for climate and land cover change impact studies. We used virtual experiments with synthetic data because they allow to unequivocally attribute changes in recharge to changes in climate and land cover. Importantly, we showed that precipitation characteristics and land cover have an interacting effect on simulated recharge, consistently with the global and observational analysis by Kim and Jackson (2012) of non-karst areas, which revealed interactions among vegetation and climate (in particular precipitation overall amount) in producing recharge. We found that simulated recharge is lower (and ET is higher) for forest compared to the shorter vegetation type considered, i.e. shrub, which is in line with expectations and in particular previous studies in karst areas (Ford and Williams, 2007; Williams, 1993). We demonstrated that the overall amount and daily intensity of precipitation have a positive effect on simulated recharge, i.e. the higher they are the higher the recharge. This is consistent with previous observational studies, which found that heavy precipitation events enhance groundwater recharge in non-karst areas (Taylor et al. (2013) in a semi-arid tropical region, and Owor et al. (2009) in a seasonally humid tropical region). On the other hand, the modelling study by Hartmann et al. (2017) showed that heavy precipitation events can have a negative effect on recharge simulated by the VarKarst model in some carbonate rock areas in Europe, the Middle East and Northern Africa. However, Hartmann et al. (2017) analysed the sensitivity of simulated recharge using forcing data from Global Circulation Models (GCMs). Given that climate variables from GCMs show complex patterns, it might be difficult to isolate the effect of a specific climate property (e.g. heavy precipitation events) on the simulated fluxes, which could explain the differences between their findings -and ours.

~~Although p~~Precipitation patterns are more complex than the simple periodic variations we used here, and the steady state conditions may never be reached in practice. Nevertheless, we believe that performing virtual experiments similar to the ones proposed in the present study is a complementary approach to application of climate projections provided by ~~Global Circulation Models (~~GCMs~~)~~ and future land cover change scenarios (e.g. Holman et al., 2017; Hurtt et al., 2011), to ~~understand~~

systematically assess the sensitivity of a model to changes in input characteristics, to test the soundness of the model sensitivities and to determine which aspects of a model input would be worth further investigating.

6.3 Applying V2Karst over larger domains

~~In this section, we first discuss the importance for large-scale applications of the new processes that we introduced in V2Karst and second the strategy to estimate the model parameters over large domains.~~

The results of our global sensitivity analyses suggest that all newly introduced processes into V2Karst ~~(transpiration, soil evaporation, evaporation from canopy interception, vegetation seasonality and contribution of the water stored below the root zone to transpiration)~~ are relevant for applications over large domains because ~~all~~ the parameters that control these processes ~~of them~~ can affect simulated recharge, depending on the climatic, soil and land cover conditions. Specifically, ~~the results of our sensitivity analyses across a large range of soil and land cover conditions (wide unconstrained ranges) showed that overall~~ transpiration and vegetation seasonality are important processes under the climate of the four FLUXNET sites, ~~and additionally~~ evaporation from canopy interception at all forested sites (German site and two French sites), ~~and~~ the contribution of water stored below the root zone ~~are also important model components~~ under the climate of the ~~German site and~~ Spanish site ~~respectively. Moreover, the sensitivity analysis using site-specific constrained ranges revealed that the process of evaporation from canopy interception has an effect on simulated recharge at all forested sites (German site and two French sites)~~ and ~~that the process of~~ soil evaporation ~~has an impact on simulated recharge~~ at the semi-arid site with sparse and short vegetation (Spanish site). The importance of representing canopy interception, in particular for forested land covers, was already mentioned in previous studies (Gerrits, 2010; Savenije, 2004) and the significance of separating transpiration and soil evaporation was reported in Maxwell and Condon (2016) and Wang and Dickinson (2012).

Regarding parameter estimation, the use of soft rules similar to this study and to the large-scale study of Hartmann et al. (2015) can be envisaged for applications over large domains. We showed that a priori information on parameter ranges is needed to constrain the simulations. At large-scales, a priori information on vegetation parameters can be derived from large-scale databases of vegetation properties (more details in Sect. S1 of our Supplementary material) and a priori information on parameters that control the subsurface properties (V_{soi}) can be derived following Hartmann et al. (2015). It is also particularly important to assess the sensitivity of model parameters across the modelling domain to test the suitability of fixing model parameters, as done in this study at FLUXNET sites. ~~w~~We indeed found that parameters V_{can} and $r_{s,soi}$, that are typically fixed in the other large-scale models of Table A1, do have an impact on ~~total~~ simulated recharge at least one FLUXNET site. Moreover, as mentioned in Sect. 2.3, V_{can} and $r_{s,soi}$ are understood to vary across land cover type and soil type respectively, even if no clear ranges of these parameters have been established across land cover and soil types respectively. Therefore, fixing these two parameters could potentially introduce large uncertainties in V2Karst simulations. Other studies have reported on the issue, and in particular Cuntz et al. (2016) showed that some constant parameters of the Noah-MP land surface model can be highly influential for some model outputs. Finally, parameter estimation for application over large domains will need to account for the large variability in hydraulic properties and in recharge patterns observed across karst systems reported e.g. in Hartmann et al. (2014) and in Klimchouk and Ford (2000), for instance using a simplified classification of karst landscapes based on climate and topography as in Hartmann et al. (2015).

We note that although soils may be absent in some karst areas (e.g. karren field in high mountain areas as reported in Hartmann et al., 2014), V2Karst always includes a soil layer. However, this soil layer can be very thin and have a limited impact on recharge. Additionally, when the simulation units have a large extent, such as 0.25°x0.25° in Hartmann et al. (2015), it is reasonable to assume that soil layers can always be found.

The model can account for sub-grid heterogeneity in vegetation type using a 'tile' approach as in Mac-PDM (Gosling and Arnell, 2011) and VIC (Bohn and Vivoni, 2016), which consists of subdividing each model grid cell in a number of independent units (tiles), each of which has a specific land cover (e.g. short or tall vegetation). The model is then evaluated separately over each tile and the overall simulated fluxes are computed as the area weighted average of the fluxes calculated over the tiles. In V2Karst, each 'tile' includes a vegetated and non-vegetated fraction as explained in Sect. 2.3.3.

~~Regarding the estimation of V2Karst parameters, in this study, we showed that the application of the soft rules based on the comparison between observed and simulated variables and on a priori information on parameter ranges (Sect. 4.1) allowed to estimate V2Karst parameters and constrain the model predictions at the four FLUXNET sites. Therefore, to confine V2Karst parameter ranges over a large modelling domain, future studies will investigate the application of an approach similar to the strategy presented in this study and in (Hartmann et al., 2015) for the VarKarst model, based on soft rules and on the grouping of the model grid cells across the application domain into typical karst-vegetation landscapes. In addition to a priori information on the value of the soil water capacity V_{soi} used in (Hartmann et al., 2015), a priori information on the vegetation parameters will also need to be derived from large-scale databases of vegetation properties (more details on these databases in Sect. S1 of our supplementary material). We can anticipate that the estimation of the parameters that characterise sub-surface properties (a, V_{soi}, V_r) may be particularly critical, since our sensitivity analyses using site-specific constrained parameter ranges showed that these parameters have the largest impact on simulated recharge. In addition, unlike above-ground vegetation properties that can be more easily observed (e.g. LAI, vegetation height), sub-surface properties are not often well investigated.~~

~~One question that we think is still insufficiently addressed in large-scale hydrological modelling is the issue of which parameters should be varied during parameter estimation and uncertainty analysis, and instead which parameters can be reasonably fixed to a constant value across the modelling domain to simplify the analyses. Other studies have reported on the issue, and in particular a study by (Cuntz et al., 2016) showed that some constant parameters of the Noah-MP land surface model can be highly influential for some model outputs. Likewise, in this study, we found that parameters V_{can} and $r_{s,soi}$, that are typically fixed in the other large-scale models of Table A1, do have an impact on total recharge at least one FLUXNET~~

~~site. Moreover, as mentioned in Sect. 2.3, V_{can} and $r_{s,soi}$ are understood to vary across land cover type and soil type respectively, even if no clear ranges of these parameters have been established across land cover and soil types respectively. Therefore, fixing these two parameters could potentially introduce large uncertainties in V2Karst simulations.~~

~~The reason for the modellers' decision to fix a given parameter could for example have been based on the finding that the parameter might not have been influential for a particular site at which sensitivity was analysed. However, it might be that the same parameter is influential for other systems with different characteristics since parameter sensitivity can show a high variability across places as suggested by this study and as further demonstrated in (Güntner et al., 2007; Van Werkhoven et al., 2008a). It is therefore particularly important to assess the sensitivity of model parameters across the modelling domain to test the suitability of fixing model parameters, as done in this study at FLUXNET sites.~~

7. Conclusions

The objectives of the present study were (1) to develop and test an ET component with explicit representation of land cover processes for the ~~large-scale~~ karst recharge model VarKarst, so that the model can be used for combined climate and land cover change impact assessment at large-scales, and (2) to evaluate the mechanisms of recharge production in the model as well as the model sensitivity to temporal precipitation patterns and land cover using ~~virtual experiment~~observations and synthetic data to force the model.

Many different approaches are used to represent ET in large-scale hydrologic models, and the lack of in-situ ET observations makes it difficult to assess and compare the performance of these different formulations. Moreover, some models use a large number of parameters that can be only poorly constrained by the few available observations. High model complexity also makes Monte Carlo simulation computationally expensive and hampers uncertainty and sensitivity analysis. The new version of the VarKarst model developed here, V2Karst (V1.0), is the first large-scale hydrological model to include explicit representations of both karst and land cover processes. We sought to include parsimonious process ~~es~~descriptions that are understood to be relevant for climate and land cover impact assessment, namely, (1) a representation of the three ET components (transpiration, soil evaporation in presence of sparse canopy and evaporation from canopy interception) and (2) a physically-based PET equation (Penman-Monteith). The model also comprises a parsimonious representation of vegetation seasonality.

We ~~showed~~ demonstrated that V2Karst ~~was able to~~ produces ~~plausible~~credible simulations at four carbonate rock FLUXNET sites, ~~since~~ by showing that ~~its simulations were consistent with~~ it reproduces observations of latent heat and soil moisture ~~and a priori information at the sites,~~ and that the parameters that dominate the model sensitivity ~~were~~ are in accordance with our perception of expected controls on recharge. ~~Additionally, it was also shown that all newly introduced processes in V2Karst can have an impact on simulated recharge depending on the climate, the soil properties and the land cover.~~We also established that the model has plausible sensitivities to climate and land cover changes when using virtual experiments with synthetic precipitation and land cover scenarios.

Additionally, ~~it was also shown that all newly~~we showed that the newly introduced processes in V2Karst can have varying~~an~~ impacts on simulated recharge quantities depending on the climate, the soil properties and the land cover.

Overall, our study demonstrates the value of a model development and evaluation process that considers both how well a model reproduces historical observations as well as how this performance is achieved. The latter is examined by using global sensitivity analysis to understand which processes and inputs dominate the model output. Through the proposed process we have a higher chance of obtaining the right results for the right reasons than by using performance analysis alone (Kirchner, 2006). We further believe– given the lack of observations of many hydrological fluxes at large scales – that an integrated analysis using historical data, our expectations derived from previous studies and synthetic data (to understand model behaviour more generally) can provide a more holistic view of a model than using observations alone.

~~Virtual experiments, using synthetic periodic precipitation inputs to force the model, allowed to characterise the sensitivity of simulated recharge to the precipitation temporal distribution, the precipitation amount, the seasonal conditions of the other climate variables and the land cover. This had been little examined in previous studies in karst areas. Our results call for a large-scale assessment of the combined impact of future changes in climate (and more specifically the precipitation amount 10 and temporal distribution) and in land cover on groundwater recharge in karst areas.~~

~~Importantly, our study demonstrate that global sensitivity analysis can provide valuable insights for model development, since it can help to determine which processes should be included in models and which parameters can be fixed to constant values with little impact on the simulations. Moreover, global sensitivity analysis, allows to characterise a model sensitivity to changes in climate and land cover. We therefore believe that large-scale hydrology would benefit from a more exhaustive evaluation of the models' sensitivities over their application domain, since so far sensitivity analyses of large-scale models are very few and many of them explore a limited ranges of possible parameter combinations only.~~

Appendix A. Review of ET component in large-scale hydrological models

Model	Δt	Sub-grid variability of soil moisture [a]	Energy balance	ET processes					Number of parameters for ET estimation [b]	Reference
				T_{act}^{over}	T_{act}^{under}	Es_{act}	Ec_{act}	Carbon cycle		
WBM	daily	no	no	**yes**	no	no	no	no	3 (minimum)	(Federer et al., 2003; Vörösmarty et al., 1989; Vörösmarty et al., 1998)
~~LaD~~	~~sub-daily~~	~~no~~	~~**yes**~~	~~**yes**~~	~~no~~	~~no~~	~~no~~	~~no~~	~~5~~	~~(Milly and Shmakin, 2002)~~
WaterGap V2.2	daily	**implicit**	no	**yes**	no	no	**yes**	no	7	(Döll et al., 2003; Müller Schmied et al., 2014)
mHM	daily/sub-daily	**implicit**	no	**yes**	no	no	**yes**	no	10[c]	(Kumar et al., 2013; Samaniego et al., 2010, 2018)
LPJ	daily	no	no	**yes**	no	**yes**	**yes**	**yes**	14	(Gerten et al., 2004; Sitch et al., 2003)
Model of (Kergoat, 1998)	daily	no	no	**yes**	no	**yes**	**yes**	no	15	(Kergoat, 1998)
PCR-GLOBWB	daily	**implicit**	no	**yes**	no	**yes**	**yes**	no	15	(Van Beek and Bierkens, 2008; Van Beek, 2008; Sperna Weiland et al., 2015; Sutanudjaja et al., 2011)
Mac-PDM	daily	**implicit**	no	**yes**	**yes**	no	**yes**	no	16 [d~~e~~]	(Arnell, 1999; Gosling and Arnell, 2011; Smith, 2016)
~~ISBA~~	~~sub-daily~~	~~**implicit**~~	~~**yes**~~	~~**yes**~~	~~no~~	~~**yes**~~	~~**yes**~~	~~no~~	~~17~~	~~(Boone et al., 1999; Decharme and Douville, 2006; Noilhan and Planton, 1989)~~
GLEAM V3	daily	no	no	**yes**	no	**yes**	**tall land cover**	no	18 [e~~d~~]	(Martens et al., 2017; Miralles et al., 2010, 2011)
VIC V4.2	daily/ sub-daily	**implicit**	**optional**	**yes**	no	**yes**	**yes**	no	22	(Bohn and Vivoni, 2016; Liang et al., 1994)

Table A1. Characteristics of selected large-scale models: simulation time step (Δt), representation of sub-grid variability of soil moisture, solving of the energy balance, ET processes represented (Overstory transpiration T_{act}^{over}, understory transpiration T_{act}^{under}, soil evaporation Es_{act}, evaporation from canopy interception Ec_{act}, and carbon cycle i.e. vegetation dynamic model), and number of parameters for ET estimation. The models were selected based on the following criteria: (1) explicit representation of land cover properties, (2) calculation of ET and soil water balance at a daily or sub-daily time step, and (3) applications in previous studies over a wide range of climate and land cover types. Tables A2 and A3 present the parameterisations used in these models.

[a] None of these models account for karst processes as done by the VarKarst model (Hartmann et al., 2015).
[b] Number of parameters for a given land cover type, excluding parameters used in the representation of vegetation seasonality, carbon cycle (vegetation dynamic), sublimation from snowpack and snowmelt evaporation to make models more comparable.

[c] Number of parameters considering three soil layers.
[d]~~[e]~~ This number includes the parameters used for the computation of both understory and overstory (grass) transpiration.
[e]~~[d]~~ Number of parameters assuming tall vegetation (interception is considered for tall vegetation only).

Model	Potential evapotranspiration (PET)			Stress model for actual ET calculation from PET
	Formulation [a]	Surface resistance r_s	Number of parameters	
WBM	T, SW	constant when considered	0 (minimum)	function of soil moisture which multiplies PET
~~LaD~~	~~PM~~	~~constant~~	~~2~~	~~function of soil moisture which multiplies PET~~
WaterGap V2.2	PT	not included	1	demand-supply model (Federer, 1982)
mHM	HS	Not included	1 (aspect correction)	function of soil moisture which multiplies PET
LPJ	empirical formula based on PT	function of CO_2 and photosynthesis	4	demand-supply model for transpiration (Federer, 1982) and function of soil moisture which multiplies PET for soil evaporation
Model of (Kergoat, 1998)	PM	Jarvis type (Jarvis, 1976; Stewart, 1988)	10	function of soil moisture which multiplies r_s
PCR-GLOBWB	PM	empirical reference crop scheme (Allen et al., 1998) [b]	2	function of soil moisture and soil hydraulic properties which multiplies PET
Mac-PDM	PM	constant	8 [c]	function of soil moisture which multiplies PET
~~ISBA~~	~~PM~~	~~Jarvis type (Jarvis, 1976; Stewart, 1988)~~	~~8~~	~~function of soil moisture which multiplies r_s for transpiration and PET for soil evaporation~~
GLEAM V3	PT	not included	3 [d]	function of soil moisture and vegetation optical depth which multiplies PET
VIC V4.2	PM	Jarvis type (Jarvis, 1976; Stewart, 1988)	12	function of soil moisture which multiplies r_s for transpiration and PET for soil evaporation

Table A2. Representation of potential evapotranspiration (PET) and stress model for actual evapotranspiration (ET) calculation from PET in the large-scale models of Table A1.

[a] T: Thornthwaite (1948); HS: Hargreaves and Samani (1985); PT: Priestley-Taylor (1972); PM: Penman-Monteith (Monteith, 1965); SW: Shuttleworth and ~~-~~Wallace (1985).

[b] This approach consists of calculating a value of PET for a reference grass surface with known properties and to adjust this potential rate using land cover specific empirical crop factors. This formulation avoids the specification of the stomatal resistance whose value is largely uncertain (see Sect. S1 of our supplementary material). Tabulated values of the crop factors for agricultural crops are provided in (Allen et al., 1998). However, the origin of the crop factor formulation for non-agricultural crops is not clear.

[c] This number includes the parameters used for the computation of PET for both understory and overstory (grass).

[d] Number of parameters assuming tall vegetation (interception is considered for tall vegetation only).

Model	Sparse vegetation formulation [a]	Soil layers [b]	Evaporation from canopy interception (Ec_{act})		Seasonality of vegetation
			model	Number of parameters	
WBM	not included	1 layer	not included	0	not included
~~LaD~~	~~not included~~	~~1 layer~~	~~not included~~	~~0~~	~~not included~~
WaterGap V2.2	not included	1 layer	overflow store	3	empirical LAI growth model
mHM	Not included	3 layers (T_{act} from all layers depending on their relative root fractions)	overflow store	2	monthy values of canopy interception capacity calculated from monthly LAI
LPJ	uncoupled (vegetated and bare soil tiles)	3 layers (Es_{act} from shallow layer and T_{act} from all layers depending on their relative root fractions)	empirical: fraction of precipitation (Kergoat, 1998)	2	vegetation dynamic model
Model of (Kergoat, 1998)	coupled moisture uptake	1 layer	empirical: fraction of precipitation (Kergoat, 1998)	2	LAI set to zero during leaf-off season
PCR-GLOBWB	coupled moisture uptake	2 layers (Es_{act} from shallow layer and T_{act} from all layers depending on their relative root fractions)	overflow store	2	monthly values of crop factors and LAI
Mac-PDM	uncoupled (overstory and understory tiles)	1 layer for each tile	Calder (Calder, 1990)	3 [c]	not included
~~ISBA~~	~~coupled moisture uptake~~	~~3 layers (Es_{act} from two shallower layers and T_{act} from middle layer and capillary rise from deeper layer)~~	~~overflow store~~	~~3~~	~~monthly values of vegetation parameters~~
GLEAM V3	uncoupled (vegetated and bare soil tiles)	3 layers (Es_{act} from shallower layer and T_{act} in wettest layer)	Gash (Gash, 1979; Valente et al., 1997)	7 [d]	assimilation of vegetation optical depth
VIC V4.2	coupled moisture uptake	2 layers (Es_{act} from shallower layer and T_{act} from all layers depending on their relative root fractions)	overflow store	3	monthly values of LAI and assimilation of daily NDVI

Table A3. Representation of sparse vegetation, soil layers, evaporation from canopy interception and seasonality of vegetation in the large-scale models of Table A1.

[a] Uncoupled approaches consist of assessing separately the water balance for the vegetated and bare soil fractions (overstory and understory fractions for Mac-PDM). Therefore, this approach is based on the simplifying assumption that the vegetation roots do not extent beyond the surface area covered by the vegetation canopy. Instead, coupled approaches evaluate the overall water balance over both fraction, thus allowing for interactions for soil moisture uptake between vegetated and bare soil fractions. All models neglect aerodynamic interactions between vegetation and bare soil. This can be accounted for using for instance the Shuttleworth-Wallace PET equation (Shuttleworth and Wallace, 1985), which requires the specification of further resistance parameters compared to the Penman-Monteith equation. The Shuttleworth-Wallace equation was used anecdotally in the WBM model for a few applications.
[b] Es_{act}: actual soil evaporation; T_{act}: actual vegetation transpiration.
[c] This number includes the parameters used for the computation of PET for both understory and overstory (grass).
[d] Number of parameters assuming tall vegetation (interception is considered for tall vegetation only).

Appendix B. Additional information on the four carbonate rock FLUXNET sites

Site name		Hainich (German site)	Llano de los Juanes (Spanish site)	~~Puéchabon (French 1 site)~~Font-Blanche (French 2 site)	~~Font-Blanche (French 2 site)~~Puéchabon (French 1 site)
General information	Coordinates	51°04′45″N, 10°27′07″E	36°55’56’’N, 2°44’55’’W	43°14′27″N, 5°40′45″E	43°44’29’’N, 3°35’45’’E
	Elevation	430 m a.s.l.	1600 m a.s.l	420 m a.s.l	270 m a.s.l
Vegetation	Type	Deciduous broadleaf trees	Shrubs, herbs, bare soil, rock outcrops	Evergreen trees (30% broadleaf and 70% needleleaf)	Evergreen broadleaf trees
	Maximum *LAI*	5 $m^2.m^{-2}$	2.71 $m^2.m^{-2}$	2.2 $m^2.m^{-2}$	2.9 ± 0.4 $m^2.m^{-2}$
	Height	Around 33 m	0.5 m (average) - 1.2 m (maximum)	6 m (broadleaf) and 12 m (needleleaf)	5.5 m
	Seasonality	Leaves from May to October	1.31 $m^2.m^{-2}$ (annual minimum)	Not available	Not available
	Rooting depth	Not available	Roots probably access water below the soil	Roots probably access water below the soil	4.5 m (150 mm available water capacity)
Soil	Texture	Silty clay	Silt loam and clay loam	Sandy clay loam	Silty clay loam and clay loam
	Depth	0.5 - 0.7 m	0.1 – 0.3 m (occasionally up to 1.5 m)	0.6 m (maximum)	No clear limit between soil and epikarst
	Available water capacity [a]	0.13 $m^3.m^{-3}$	0.25 $m^3.m^{-3}$	49 mm	No clear limit between soil and epikarst
	Other properties	Permeable loess layer of 10 -50 cm between soil and bedrock	Rocky soil	Rocky soil	Rocky soil
Bedrock		Fissured and fractured limestone	Karstified dolomite and dolines	Karstified limestone	Karstified limestone
Hydrology	Surface runoff	Low	Low	Low	Inexistent
	Recharge	Large part of the water balance	Diffuse and concentrated, high temporal variability	Not available	Not available
Measure-ments	Height for humidity and temperature	43.5 m	1.5m	16 m	12.2 m
	Height for wind speed	43.5 m	2.5 m	16 m	12.2 m
	Depth for soil moisture	0.05, 0.15, 0.3 m	0.15 m	Not measured	Not measured
References		(Knohl et al., 2003; Mund et al., 2010; Pinty et al., 2011), personal communication from	(Alcalá et al., 2011; Cantón et al., 2010; Contreras et al., 2008; Li et al., 2007, 2011; Pérez-	(Ecofor, n.d.; Gea-Izquierdo et al., 2015; Simioni et al., 2013), personal	(Rambal, 1992, 2011; Rambal et al., 2003; Reichstein et al., 2002)

	Martina Mund and Manfred Fink	Priego et al., 2013; Serrano-Ortiz et al., 2007)	communication from Guillaume Simioni,

Table B1. Description of the four carbonate rock FLUXNET sites. [a] between wilting point and field capacity.

Supplementary material

Code availability

The code of the V2Karst model is open source and freely available under the terms of the GNU General Public License version 3.0. The model code is written in matlab and is provided through a Github repository: https://github.com/fannysarrazin/V2Karst_model

Author contribution

F.S., A.H., R.R and T.W. ~~contributed to the~~ designed ~~of~~ the model equations of V2Karst. F.S., A.H., F.P. and T.W. contributed to the methodology to test the model. ~~and the analysis of the results.~~ F.S. led the analyses, performed the numerical experiments and the processing of the data, and developed the code of the V2Karst model building on the code of the previous version of the model (VarKarst), which was developed by A.H.. ~~F.S. performed the numerical experiments. F.S., A.H., F.P. and T.W. contributed to the methodology and the analysis of the results.~~ A.H. contributed more particularly to the karst aspect of the work, R.R. to the land surface modelling aspect of the work and to the processing of the data, and F.P and T.W. to the sensitivity analysis aspect of the work. The manuscript was prepared by F.S. and edited by all the authors.

Competing interest

The authors declare that they have no conflict of interest.

Acknowledgements

F.S. was supported by a PhD Scholarship Programme from the University of Bristol Development and Alumni Relations Office. Support to A.H. was provided by the Emmy Noether-Programme of the German Research Foundation (DFG; grant number HA 8113/1-1; project ‘Global Assessment of Water Stress in Karst Regions in a Changing World’). F.P. was partially supported by a UK Engineering and Physical Sciences Research Council fellowship (EPSRC; grant number EP/R007330/1). ~~We are grateful to Rafael Rosolem for advice on FLUXNET data processing and ET modelling.~~ Support to R.R. was provided by the NERC ‘A MUlti-scale Soil moisture Evapotranspiration Dynamics study’ (AMUSED project; grant number NE/M003086/1) and the ‘Brazilian Experimental datasets for MUlti-Scale interactions in the critical zone under Extreme Drought’ (BEMUSED project, grant number NE/R004897/1)~~, both projects funded by Natural Environment Research Council (NERC)~~. T.W. was partially supported by a Royal Society Wolfson Research Merit Award. We thank Timothy Foster for advice on ET modelling. We thank Yoshihide Wada for providing information on ET representation in the PCR-GLOBWB

model. We thank the investigators of the four FLUXNET sites (DE-Hai, Es-Lju, FR-FBn and FR-Pue) for allowing us to use their data, and in particular Martina Mund and Manfred Fink (De-Hai) Penelope Serrano Ortiz, Francisco Domingo Poveda and Andrew Kowalski (ES-Lju), Richard Joffre and Serge Rambal (FR-Pue), and Guillaume Simioni (FR-FBn) for sharing information on the characteristics of the FLUXNET sites. We appreciate the support of Eleonora Canfora and Dario Papale from the FLUXNET European Fluxes Database Cluster.

References

Abramowitz, G., Leuning, R., Clark, M. and Pitman, A.: Evaluating the performance of land surface Models, J. Clim., 21(21), 5468–5481, doi:10.1175/2008JCLI2378.1, 2008.

Alcalá, F. J., Cantón, Y., Contreras, S., Were, A., Serrano-Ortiz, P., Puigdefábregas, J., Solé-Benet, A., Custodio, E. and Domingo, F.: Diffuse and concentrated recharge evaluation using physical and tracer techniques: Alcalá, Francisco J., Yolanda Cantón, Sergio Contreras, Ana Were, Penélope Serrano-Ortiz, Juan Puigdefábregas, Albert Solé-Benet, Emilio Custodio, and Francisco Domingo. 20, Environ. Earth Sci., 62(3), 541–557, doi:10.1007/s12665-010-0546-y, 2011.

Allen, R. G., Pereira, L. S., Raes, D. and Smith, M.: Crop evapotranspiration: Guidelines for computing crop requirements, FAO Irrigation and Drainage Paper 56, Food and Agriculture Organization (FAO), Rome, Italy., 1998.

Allen, R. G., Pruitt, W. O., Wright, J. L., Howell, T. A., Ventura, F., Snyder, R., Itenfisu, D., Steduto, P., Berengena, J. and Yrisarry, J. B.: A recommendation on standardized surface resistance for hourly calculation of reference ETo by the FAO56 Penman-Monteith method, Agric. Water Manag., 81(1–2), 1–22, doi:10.1016/j.agwat.2005.03.007, 2006.

Arbel, Y., Greenbaum, N., Lange, J. and Inbar, M.: Infiltration processes and flow rates in developed karst vadose zone using tracers in cave drips, Earth Surf. Process. Landforms, 35(14), 1682–1693, doi:10.1002/esp.2010, 2010.

Archfield, S. A., Clark, M., Arheimer, B., Hay, L. E., McMillan, H., Kiang, J. E., Seibert, J., Hakala, K., Bock, A., Wagener, T., Farmer, W. H., Andréassian, V., Attinger, S., Viglione, A., Knight, R., Markstrom, S. and Over, T.: Accelerating advances in continental domain hydrologic modeling, Water Resour. Res., 51(12), 10078–10091, doi:10.1002/2015WR017498, 2015.

Arnell, N. W.: A simple water balance model for the simulation of streamflow over a large geographic domain, J. Hydrol., 217(3–4), 314–335, doi:10.1016/S0022-1694(99)00023-2, 1999.

Atkinson, S. E., Woods, R. A. and Sivapalan, M.: Climate and landscape controls on water balance model complexity over changing timescales, Water Resour. Res., 38(12), 50-1-50–17, doi:10.1029/2002WR001487, 2002.

Bai, P., Liu, X., Liang, K. and Liu, C.: Comparison of performance of twelve monthly water balance models in different climatic catchments of China, J. Hydrol., 529, 1030–1040, doi:10.1016/j.jhydrol.2015.09.015, 2015.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K. and Wofsy, S.: FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem-Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities, Bull. Am. Meteorol. Soc., 82(11), 2415–2434, doi:10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2, 2001.

Bargués Tobella, A., Reese, H., Almaw, A., Bayala, J., Malmer, A., Laudon, H. and Ilstedt, U.: The effect of trees on preferential flow and soil infiltrability in an agroforestry parkland in semiarid Burkina Faso, Water Resour. Res., 50(4), 3342–3354, doi:10.1002/2013WR015197, 2014.

Beven, K. and Germann, P.: Macropores and water flow in soils revisited, Water Resour. Res., 49(6), 3071–3092, doi:10.1002/wrcr.20156, 2013.

Beven, K. J. and Cloke, H. L.: Comment on “hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth’s terrestrial water” by Eric F. Wood et al., Water Resour. Res., 48(1), W01801, doi:10.1029/2010WR010090, 2012.

Bierkens, M. F. P.: Global hydrology 2015: State, trends, and directions, Water Resour. Res., 51(7), 4923–4947, doi:10.1002/2015WR017173, 2015.

Blume, H.-P., Brümmer, G. W., Horn, R., Kandeler, E., Kögel-Knabner, I., Kretzschmar, R., Stahr, K. and Wilke, B.-M.: Lehrbuch der Bodenkunde, Springer-Verlag, Berlin Heidelberg, doi: 10.1007/978-3-662-49960-3., 2010.

Bohn, T. J. and Vivoni, E. R.: Process-based characterization of evapotranspiration sources over the North American monsoon region, Water Resour. Res., 52(1), 358–384, doi:10.1002/2015WR017934, 2016.

Boone, A., Calvet, J.-C. and Noilhan, J.: Inclusion of a third soil layer in a land surface scheme using the force–restore method, J. Appl. Meteorol., 38(11), 1611–1630, doi:10.1175/1520-0450(1999)038<1611:IOATSL>2.0.CO;2, 1999.

Botter, G., Porporato, A., Rodriguez-Iturbe, I. and Rinaldo, A.: Nonlinear storage-discharge relations and catchment streamflow regimes, Water Resour. Res., 45(10), W10427, doi:10.1029/2008WR007658, 2009.

Brown, A. E., Zhang, L., Mcmahon, T. A., Western, A. W. and Vertessy, R. A.: A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation, J. Hydrol., 310(1–4), 28–61, doi:10.1016/j.jhydrol.2004.12.010, 2005.

Calder, I. R.: Evaporation in the Uplands, John Wiley & Sons, Ltd, Chichester, UK., 1990.

Campolongo, F., Cariboni, J. and Saltelli, A.: An effective screening design for sensitivity analysis of large models, Environ. Model. Softw., 22(10), 1509–1518, doi:10.1016/j.envsoft.2006.10.004, 2007.

Campolongo, F., Saltelli, A. and Cariboni, J.: From screening to quantitative sensitivity analysis. A unified approach, Comput. Phys. Commun., 182(4), 978–988, doi:10.1016/j.cpc.2010.12.039, 2011.

Canora, F., Fidelibus, M. D., Sciortino, A. and Spilotro, G.: Variation of infiltration rate through karstic surfaces due to land use changes: A case study in Murgia (SE-Italy), Eng. Geol., 99(3–4), 210–227, doi:10.1016/j.enggeo.2007.11.018, 2008.

Cantón, Y., Villagarcía, L., José Moro, M., Serrano-Ortíz, P., Were, A., Javier Alcalá, F., Kowalski, A. S., Solé-Benet, A., Lázaro, R. and Domingo, F.: Temporal dynamics of soil water balance components in a karst range in southeastern Spain: estimation of potential recharge, Hydrol. Sci. J., 55(5), 737–753, doi:10.1080/02626667.2010.490530, 2010.

Chaney, N. W., Herman, J. D., Ek, M. B. and Wood, E. F.: Deriving global parameter estimates for the Noah land surface model using FLUXNET and machine learning, J. Geophys. Res., 121(22), 13,218-13,235, doi:10.1002/2016JD024821, 2016.

Chen, Z., Auler, A. S., Bakalowicz, M., Drew, D., Griger, F., Hartmann, J., Jiang, G., Moosdorf, N., Richts, A., Stevanovic, Z., Veni, G. and Goldscheider, N.: The World Karst Aquifer Mapping project: concept, mapping procedure and map of Europe, Hydrogeol. J., 25(3), 771–785, doi:10.1007/s10040-016-1519-3, 2017.

Clapp, R. B. and Hornberger, G. M.: Empirical equations for some soil hydraulic properties, Water Resour. Res., 14(4), 601–604, doi:10.1029/WR014i004p00601, 1978.

Contreras, S., Boer, M. M., Alcala, F. J., Domingo, F., Garcia, M., Pulido-Bosch, A. and Puigdefabregas, J.: An ecohydrological modelling approach for assessing long-term recharge rates in semiarid karstic landscapes, J. Hydrol., 351(1–2), 42–57, doi:10.1016/j.jhydrol.2007.11.039, 2008.

COST: Cost action 65 - Hydrogeological aspects of groundwater protection in karstic areas, Report EUR 16547, European Commission, Directorate-General XII Science, Research Development, Luxembourg., 1995.

Coxon, C.: Agriculture and Karst, in Karst Management, edited by P. E. van Beynen, Springer Netherlands, Dordrecht., 103–138, 2011.

Cuntz, M., Mai, J., Samaniego, L., Clark, M., Wulfmeyer, V., Branch, O., Attinger, S. and Thober, S.: The impact of standard and hard-coded parameters on the hydrologic fluxes in the Noah-MP land surface model, J. Geophys. Res., 121(18), 10,676-10,700, doi:10.1002/2016JD025097, 2016.

Cuthbert, M. O., Mackay, R. and Nimmo, J. R.: Linking soil moisture balance and source-responsive models to estimate diffuse and preferential components of groundwater recharge, Hydrol. Earth Syst. Sci., 17, 1003–1019, doi:10.5194/hess-17-1003-2013, 2013.

DeFries, R. and Eshleman, K. N.: Land-use change and hydrologic processes : a major focus for the future, Hydrol. Process., 18(11), 2183–2186, doi:10.1002/hyp.5584, 2004.

De Groen, M. M.: Modelling interception and transpiration at monthly time steps : introducing daily variability through Markov chains, Ph.D. thesis, Delft University of Technology, Delft, The Netherlands. [online] Available from: https://books.google.com/books?id=-birGVG44nkC&pgis=1, 2002.

Döll, P. and Fiedler, K.: Global-scale modeling of groundwater recharge, Hydrol. Earth Syst. Sci., 12(3), 863–885, doi:10.5194/hess-12-863-2008, 2008.

Döll, P., Kaspar, F. and Lehner, B.: A global hydrological model for deriving water availability indicators: Model tuning and validation, J. Hydrol., 270(1–2), 105–134, doi:10.1016/S0022-1694(02)00283-4, 2003.

Doummar, J., Sauter, M. and Geyer, T.: Simulation of flow processes in a large scale karst system with an integrated catchment model (Mike She) – Identification of relevant parameters influencing spring discharge, J. Hydrol., 426–427, 112–123, doi:10.1016/j.jhydrol.2012.01.021, 2012.

Ecofor: Site atelier de Font Blanche, [online] Available from: http://www.gip-ecofor.org/f-ore-t/fontBlanche.php (Accessed 13 December 2016), n.d.

Falkenmark, M. and Rockström, J.: The new blue and green water paradigm : breaking new ground for water resources planning and management, J. Water Resour. Plan. Manag., 132(3), 129–132, doi:10.1061/(ASCE)0733-9496(2006)132:3(129), 2006.

Federer, C. A.: Transpirational Supply and Demand: plant, soil, and atmospheric effects evaluated by simulation, Water Resour. Res., 18(2), 355–362, doi:10.1029/WR018i002p00355, 1982.

Federer, C. A., Vörösmarty, C. and Fekete, B.: Sensitivity of Annual Evaporation to Soil and Root Properties in Two Models of Contrasting Complexity, J. Hydrometeorol., 4(6), 1276–1290, doi:10.1175/1525-7541(2003)004<1276:SOAETS>2.0.CO;2, 2003.

Fleury, P., Plagnes, V. and Bakalowicz, M.: Modelling of the functioning of karst aquifers with a reservoir model : Application to Fontaine de Vaucluse (South of France), J. Hydrol., 345(1–2), 38–49, doi:doi.org/10.1016/j.jhydrol.2007.07.014, 2007.

Foken, T., Leuning, R., Oncley, S. R., Mauder, M. and Aubinet, M.: Corrections and Data Quality Control, in Eddy Covariance: A Practical Guide to Measurement and Data Analysis, edited by M. Aubinet, T. Vesala, and D. Papale, Springer Netherlands, Dordrecht., 85–131, 2012.

Ford, D. and Williams, P.: Karst Hydrogeology and Geomorphology, John Wiley & Sons Ltd, Chichester, UK., 2007.

Gash, J. H. C.: An analytical model of rainfall interception by forests, Quaterly J. R. Meteorol. Soc., 105(443), 43–55, doi:10.1002/qj.49710544304, 1979.

Gea-Izquierdo, G., Guibal, F., Joffre, R., Ourcival, J. M., Simioni, G. and Guiot, J.: Modelling the climatic drivers determining photosynthesis and carbon allocation in evergreen Mediterranean forests using multiproxy long time series, Biogeosciences, 12(12), 3695–3712, doi:10.5194/bg-12-3695-2015, 2015.

Gerrits, M.: The role of interception in the hydrological cycle, Ph.D. thesis, Delft University of Technology, Delft, The Netherlands. [online] Available from: https://repository.tudelft.nl/islandora/object/uuid:7dd2523b-2169-4e7e-992c-365d2294d02e?collection=research, 2010.

Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W. and Sitch, S.: Terrestrial vegetation and water balance - Hydrological evaluation of a dynamic global vegetation model, J. Hydrol., 286(1–4), 249–270, doi:10.1016/j.jhydrol.2003.09.029, 2004.

Gosling, S. N. and Arnell, N. W.: Simulating current global river runoff with a global hydrological model: Model revisions, validation, and sensitivity analysis, Hydrol. Process., 25(7), 1129–1145, doi:10.1002/hyp.7727, 2011.

Güntner, A., Stuck, J., Werth, S., Döll, P., Verzano, K. and Merz, B.: A global analysis of temporal and spatial variations in continental water storage, Water Resour. Res., 43(5), W05416, doi:10.1029/2006WR005247, 2007.

Hao, Y., Yeh, T. C. J., Gao, Z., Wang, Y. and Zhao, Y.: A gray system model for studying the response to climatic change: The Liulin karst springs, China, J. Hydrol., 328(3–4), 668–676, doi:10.1016/j.jhydrol.2006.01.022, 2006.

Hargreaves, G. H. and Samani, Z. A.: Reference crop evapotranspiration from temperature, Appl. Eng. Agric., 1(2), 96–99, doi:10.13031/2013.26773, 1985.

Hartmann, A. and Baker, A.: Modelling karst vadose zone hydrology and its relevance for paleoclimate reconstruction, Earth-Sciences Rev., 172, 178–192, doi:10.1016/j.earscirev.2017.08.001, 2017.

Hartmann, A., Lange, J., Vivó Aguado, À., Mizyed, N., Smiatek, G. and Kunstmann, H.: A multi-model approach for improved simulations of future water availability at a large Eastern Mediterranean karst spring, J. Hydrol., 468–469, 130–138, doi:10.1016/j.jhydrol.2012.08.024, 2012.

Hartmann, A., Barberá, J. A., Lange, J., Andreo, B. and Weiler, M.: Progress in the hydrologic simulation of time variant recharge areas of karst systems - Exemplified at a karst spring in Southern Spain, Adv. Water Resour., 54, 149–160, doi:10.1016/j.advwatres.2013.01.010, 2013.

Hartmann, A., Goldscheider, N., Wagener, T., Lange, J. and Weiler, M.: Karst water resources in a changing world: Review of hydrological modeling approaches, Rev. Geophys., 52(3), 218–242, doi:10.1002/2013RG000443, 2014.

Hartmann, A., Gleeson, T., Rosolem, R., Pianosi, F., Wada, Y. and Wagener, T.: A large-scale simulation model to assess karstic groundwater recharge over Europe and the Mediterranean, Geosci. Model Dev., 8(6), 1729–1746, doi:10.5194/gmd-8-1729-2015, 2015.

Hartmann, A., Gleeson, T., Wada, Y. and Wagener, T.: Enhanced groundwater recharge rates and altered recharge sensitivity to climate variability through subsurface heterogeneity, Proc. Natl. Acad. Sci., 114(11), 2842–2847, doi:10.1073/pnas.1614941114, 2017.

Haughton, N., Abramowitz, G. and Pitman, A. J.: On the predictability of land surface fluxes from meteorological variables, Geosci. Model Dev., 11(1), 195–212, doi:10.5194/gmd-11-195-2018, 2018.

Hendrickx, J. M. H. and Flury, M.: Uniform and Preferential Flow Mechanisms in the Vadose Zone, in Conceptual Models of Flow and Transport in the Fractured Vadose Zone, edited by NationalResearchCouncil, The National Academies Press, Washington, D.C., 149–188, 2001.

Hogue, T. S., Bastidas, L. A., Gupta, H. V. and Sorooshian, S.: Evaluating model performance and parameter behavior for varying levels of land surface model complexity, Water Resour. Res., 42(8), W08430, doi:10.1029/2005WR004440, 2006.

Holman, I. P., Brown, C., Janes, V. and Sandars, D.: Can we be certain about future land use change in Europe? A multi-scenario, integrated-assessment analysis, Agric. Syst., 151, 126–135, doi:10.1016/j.agsy.2016.12.001, 2017.

Hong, E.-M., Pachepsky, Y. A., Whelan, G. and Nicholson, T.: Simpler models in environmental studies and predictions, Crit. Rev. Environ. Sci. Technol., 47(18), 1669–1712, doi:10.1080/10643389.2017.1393264, 2017.

Hurtt, G. C., Chini, L. P., Frolking, S., Betts, R, A., Feddema, J., Fischer, G., Fisk, J. P., Hibbard, K., Houghton, R. A., Janetos, A., Jones, C. D., Kindermann, G., Kinoshita, T., Klein Goldewijk, K., Riahi, K., Shevliakova, E., Smith, S., Stehfest, E., Thomson, A., Thornton, P., van Vuuren, D. P. and Wang, Y. P.: Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands, Clim. Change, 109, 117–161, doi:10.1007/s10584-011-0153-2, 2011.

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by V. B. and P. M. M. Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, Cambridge University Press., 2013.

Ivanov, V. Y., Bras, R. L. and Curtis, D. C.: A weather generator for hydrological, ecological, and agricultural applications, Water Resour. Res., 43(10), doi:10.1029/2006WR005364, 2007.

Jarvis, P. G.: The Interpretation of the Variations in Leaf Water Potential and Stomatal Conductance Found in Canopies in the Field, Phil. Trans. R. Soc. Lond. B., 273(927), 593–610, doi:10.1098/rstb.1976.0035, 1976.

Jothityangkoon, C. and Sivapalan, M.: Framework for exploration of climatic and landscape controls on catchment water balance, with emphasis on inter-annual variability, J. Hydrol., 371(1–4), 154–168, doi:10.1016/j.jhydrol.2009.03.030, 2009.

Kergoat, L.: A model for hydrological equilibrium of leaf area index on a global scale, J. Hydrol., 212–213, 268–286, doi:10.1016/S0022-1694(98)00211-X, 1998.

Kim, J. H. and Jackson, R. B.: A Global Analysis of Groundwater Recharge for Vegetation, Climate, and Soils, Vadose Zo. J., 11(1), doi:10.2136/vzj2011.0021RA, 2012.

Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, Water Resour. Res., 42(3), W03S04, doi:10.1029/2005WR004362, 2006.

Klimchouk, A. B. and Ford, D. C.: Types of karst and evolution of hydrogeologic setting, in Speleogenesis. Evolution of Karst Aquifers, edited by A. B. Klimchouk, D. C. Ford, A. Palmer, and W. Dreybrodt, National Speleological Society, Huntsville, Alabama, U.S.A., 45–53, 2000.

Knohl, A., Schulze, E. D., Kolle, O. and Buchmann, N.: Large carbon uptake by an unmanaged 250-year-old deciduous forest in Central Germany, Agric. For. Meteorol., 118(3–4), 151–167, doi:10.1016/S0168-1923(03)00115-1, 2003.

Kumar, R., Samaniego, L. and Attinger, S.: Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations, Water Resour. Res., 49(1), 360–379, doi:10.1029/2012WR012195, 2013.

Laio, F., Porporato, A., Ridolfi, L. and Rodriguez-Iturbe, I.: On the seasonal dynamics of mean soil moisture, J. Geophys. Res., 107(D15), ACL 8-1–ACL 8-9, doi:10.1029/2001JD001252, 2002.

Li, X. Y., Contreras, S. and Solé-Benet, A.: Spatial distribution of rock fragments in dolines: A case study in a semiarid Mediterranean mountain-range (Sierra de Gádor, SE Spain), Catena, 70(3), 366–374, doi:10.1016/j.catena.2006.11.003, 2007.

Li, X. Y., Contreras, S., Solé-Benet, A., Cantón, Y., Domingo, F., Lázaro, R., Lin, H., Van Wesemael, B. and Puigdefábregas, J.: Controls of infiltration-runoff processes in Mediterranean karst rangelands in SE Spain, Catena, 86(2), 98–109, doi:10.1016/j.catena.2011.03.003, 2011.

Liang, X., Lettenmaier, D. P., Wood, E. F. and Burges, S. J.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models, J. Geophys. Res., 99(D7), 14415–14428, doi:10.1029/94JD00483, 1994.

Loáiciga, H. A., Maidment, D. R. and Valdes, J. B.: Climate-change impacts in a regional karst aquifer, Texas, USA, J. Hydrol., 227(1–4), 173–194, doi:10.1016/S0022-1694(99)00179-1, 2000.

Lu, Y., Liu, S., Weng, L., Wang, L., Li, Z. and Xu, L.: Fractal analysis of cracking in a clayey soil under freeze-thaw cycles, Eng. Geol., 208, 93–99, doi:10.1016/j.enggeo.2016.04.023, 2016.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Férnandez-Prieto, D., Beck, H. E., Dorigo, W. A. and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geosci. Model Dev., 10(5), 1903–1925, doi:10.5194/gmd-10-1903-2017, 2017.

Maxwell, R. M. and Condon, L. E.: Connections between groundwater flow and transpiration partitioning, Science (80-.)., 353(6297), 377–380, doi:10.1126/science.aaf7891, 2016.

McCabe, M. F., Ershadi, A., Jimenez, C., Miralles, D. G., Michel, D. and Wood, E. F.: The GEWEX LandFlux project:

Evaluation of model evaporation using tower-based and globally gridded forcing data, Geosci. Model Dev., 9(1), 283–305, doi:10.5194/gmd-9-283-2016, 2016.

Mendoza, P. A., Clark, M. P., Barlage, M., Rajagopalan, B., Samaniego, L., Abramowitz, G. and Gupta, H.: Are we unnecessarily constraining the agility of complex process-based models?, Water Resour. Res., 51(1), 716–728, doi:10.1002/2014WR015820, 2015.

Miralles, D. G., Gash, J. H., Holmes, T. R. H., De Jeu, R. A. M. and Dolman, A. J.: Global canopy interception from satellite observations, J. Geophys. Res., 115(D16), D16122, doi:10.1029/2009JD013530, 2010.

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A. and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, Hydrol. Earth Syst. Sci., 15(2), 453–469, doi:10.5194/hess-15-453-2011, 2011.

Monteith, J. L.: Evaporation and environment, Symp. Soc. Exp. Biol., 19, 205–234, 1965.

Morris, M. D.: Factorial sampling plans for preliminary computational experiments, Technometrics, 33(2), 161–174, doi:10.2307/1269043, 1991.

Müller Schmied, H., Eisner, S., Franz, D., Wattenbach, M., Portmann, F. T., Flörke, M. and Döll, P.: Sensitivity of simulated global-scale freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration, Hydrol. Earth Syst. Sci., 18(9), 3511–3538, doi:10.5194/hess-18-3511-2014, 2014.

Mund, M., Kutsch, W. L., Wirth, C., Kahl, T., Knohl, A., Skomarkova, M. V. and Schulze, E. D.: The influence of climate and fructification on the inter-annual variability of stem growth and net primary productivity in an old-growth, mixed beech forest, Tree Physiol., 30(6), 689–704, doi:10.1093/treephys/tpq027, 2010.

Owor, M., Taylor, R. G., Tindimugaya, C. and Mwesigwa, D.: Rainfall intensity and groundwater recharge: empirical evidence from the Upper Nile Basin, Environ. Res. Lett., 4(3), 35009, doi:10.1088/1748-9326/4/3/035009, 2009.

Pechlivanidis, I. G., McIntyre, N. and Wheater, H. S.: The significance of spatial variability of rainfall on simulated runoff: an evaluation based on the Upper Lee catchment, UK, Hydrol. Res., 48(6), nh2016038, doi:10.2166/nh.2016.038, 2016.

Penman, H. L.: The dependance of transpiration on weather and soil conditions, J. Soil Sci., 1(1), 74–89, doi:10.1111/j.1365-2389.1950.tb00720.x, 1950.

Pereira, L. S., Allen, R. G., Smith, M. and Raes, D.: Crop evapotranspiration estimation with FAO56: Past and future, Agric. Water Manag., 147, 4–20, doi:10.1016/j.agwat.2014.07.031, 2015.

Pérez-Priego, O., Serrano-Ortiz, P., Sánchez-Cañete, E. P., Domingo, F. and Kowalski, A. S.: Isolating the effect of subterranean ventilation on CO2 emissions from drylands to the atmosphere, Agric. For. Meteorol., 180, 194–202, doi:10.1016/j.agrformet.2013.06.014, 2013.

Pianosi, F., Sarrazin, F. and Wagener, T.: A Matlab toolbox for Global Sensitivity Analysis, Environ. Model. Softw., 70, 80–85, doi:10.1016/j.envsoft.2015.04.009, 2015.

Pinty, B., Jung, M., Kaminski, T., Lavergne, T., Mund, M., Plummer, S., Thomas, E. and Widlowski, J. L.: Evaluation of the JRC-TIP 0.01° products over a mid-latitude deciduous forest site, Remote Sens. Environ., 115(12), 3567–3581, doi:10.1016/j.rse.2011.08.018, 2011.

Porporato, A., Daly, E. and Rodríguez-Iturbe, I.: Soil water balance and ecosystem response to climate change, Am. Nat., 164(5), 625–632, doi:10.1086/424970, 2004.

Priestley, C. H. B. and Taylor, R. J.: On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters, Mon. Weather Rev., 100(2), 81–92, doi:10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2, 1972.

Rahman, M. and Rosolem, R.: Towards a simple representation of chalk hydrology in land surface modelling, Hydrol. Earth Syst. Sci., 21, 459–471, doi:10.5194/hess-21-459-2017, 2017.

Rambal, S.: Quercus ilex facing water stress: a functional equilibrium hypothesis, in Quercus ilex L. ecosystems: function, dynamics and management, Advances in vegetation science (AIVS, volume 13), edited by F. Romane and J. Terradas, Springer, Dordrecht, The Netherlands., 147–153, 1992.

Rambal, S.: Le Paradoxe hydrologique des écosystèmes méditerranéens sur des sols karstiques, in Numéro spécial des Annales de la Société d'Horticulture et d'Histoire Naturelle de l'Hérault, 61–67, 2011.

Rambal, S., Ourcival, J. M., Joffre, R., Mouillot, F., Nouvellon, Y., Reichstein, M. and Rocheteau, A.: Drought controls over conductance and assimilation of a Mediterranean evergreen ecosystem: Scaling from leaf to canopy, Glob. Chang. Biol., 9(12), 1813–1824, doi:10.1111/j.1365-2486.2003.00687.x, 2003.

Reichstein, M., Tenhunen, J. D., Roupsard, O., Ourcival, J. M., Rambal, S., Miglietta, F., Peressotti, A., Pecchiari, M., Tirone, G. and Valentini, R.: Severe drought effects on ecosystem CO2 and H2O fluxes at three Mediterranean evergreen sites: Revision of current hypotheses?, Glob. Chang. Biol., 8(10), 999–1017, doi:10.1046/j.1365-2486.2002.00530.x, 2002.

Rimmer, A. and Hartmann, A.: Simplified conceptual structures and analytical solutions for groundwater discharge using reservoir equations, in Water resources management and modeling, edited by P. Nayak, InTech, Kakinada, India., 217–238, 2012.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. . K., Walker, J. P., Lohmann, D. and Toll, D.: The Global Land Data Assimilation System, Bull. Am. Meteorol. Soc., 85(3), 381–394, doi:10.1175/BAMS-85-3-381, 2004.

Rosero, E., Yang, Z. L., Wagener, T., Gulden, L. E., Yatheendradas, S. and Niu, G.-Y.: Quantifying parameter sensitivity, interaction, and transferability in hydrologically enhanced versions of the Noah land surface model over transition zones during the warm season, J. Geophys. Res. Atmos., 115(D3), D03106, doi:10.1029/2009JD012035, 2010.

Rosolem, R., Gupta, H. V., Shuttleworth, W. J., Gonçalves de Gonçalves, L. G. and Zeng, X.: Towards a comprehensive approach to parameter estimation in land surface parameterization schemes, Hydrol. Process., 27(14), 2075–2097, doi:10.1002/hyp.9362, 2013.

Ross, J.: Radiative transfer in plant communities, in Vegetation and the Atmosphere, volume I Principles, edited by J. Monteith, Academic Press, London., 13–55, 1975.

Ruiz, L., Varma, M. R. R., Kumar, M. M. S., Sekhar, M., Maréchal, J.-C., Descloitres, M., Riotte, J., Kumar, S., Kumar, C. and Braun, J.-J.: Water balance modelling in a tropical watershed under deciduous forest (Mule Hole, India): Regolith matric storage buffers the groundwater recharge process, J. Hydrol., 380(3–4), 460–472, doi:10.1016/j.jhydrol.2009.11.020, 2010.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and Tarantola, S.: Global Sensitivity Analysis, The Primer, John Wiley & Sons, Ltd, Chichester, UK., 2008.

Samaniego, L., Kumar, R. and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resour. Res., 46(5), W05523, doi:10.1029/2008WR007327, 2010.

Samaniego, L., Brenner, J., Demirel, C. M., Jing, M., Kaluza, M., Kumar, R., Langenberg, B., Rakovec, O., Schäfer, D., Schrön, M., Schweppe, R. and Thober, S.: The mesoscale Hydrologic Model mHM, Documentation for version 5.9, Helmoltz Centre for Environmental Research (UFZ), Leipzig, Germany., 2018.

Samuels, R., Rimmer, A., Hartmann, A., Krichak, S. and Alpert, P.: Climate Change Impacts on Jordan River Flow: Downscaling Application from a Regional Climate Model, J. Hydrometeorol., 11(4), 860–879, doi:10.1175/2010JHM1177.1, 2010.

Sarrazin, F., Pianosi, F. and Wagener, T.: Global Sensitivity Analysis of environmental models: Convergence and validation, Environ. Model. Softw., 79, 135–152, doi:10.1016/j.envsoft.2016.02.005, 2016.

Sauter, M.: Quantification and Forecasting of Regional Groundwater Flow and Transport in a Karst Aquifer (Gallusquelle, Malm, SW. Germany). Ph.D. thesis, Tübinger Universität, Tübinger, Germany., 1992.

Savenije, H. H. G.: Determination of evaporation from a catchment water balance at a monthly time scale, Hydrol. Earth Syst.

Sci., 1(1), 93–100, doi:10.5194/hess-1-93-1997, 1997.

Savenije, H. H. G.: The importance of interception and why we should delete the term evapotranspiration from our vocabulary, Hydrol. Process., 18(8), 1507–1511, doi:10.1002/hyp.5563, 2004.

Scanlon, B. R., Keese, K. E., Flint, A. L., Flint, L. E., Gaye, C. B., Edmunds, W. M. and Simmers, I.: Global synthesis of groundwater recharge in semiarid and arid regions, Hydrol. Prpcesses, 20(15), 3335–3370, doi:doi.org/10.1002/hyp.6335, 2006.

Schwinning, S.: The ecohydrology of roots in rocks, Ecohydrology Bearings - Invited Commentary, Ecohydrology, 3(2), 238–245, doi:10.1002/eco.134, 2010.

Seidl, R., Schelhaas, M.-J., Rammer, W. and Verkerk, P. J.: Increasing forest disturbances in Europe and their impact on carbon storage, Nat. Clim. Chang., 4, 806–810, doi:10.1038/nclimate2393, 2014.

Serrano-Ortiz, P., Kowalski, A. S., Domingo, F., Rey, A., Pegoraro, E., Villagarcía, L. and Alados-Arboledas, L.: Variations in daytime net carbon and water exchange in a montane shrubland ecosystem in southeast Spain, Photosynthetica, 45(1), 30–35, doi:10.1007/s11099-007-0005-5, 2007.

Shuttleworth, W. J.: Evapotranspiration, in Handbook of Hydrology, edited by D. R. Maidment, McGraw-Hill inc., New York., 4.1-4.53, 1993.

Shuttleworth, W. J.: Terrestrial Hydrometeorology, John Wiley & Sons, Ltd, Chichester, UK., 2012.

Shuttleworth, W. J. and Wallace, J. S.: Evaporation From Spare Crops - An Energy Combination Theory, Q. J. R. Meteorol. Soc., 111(469), 839–855, doi:10.1002/qj.49711146910, 1985.

Simioni, G., Durand-Gillmann, M. and Huc, R.: Asymmetric competition increases leaf inclination effect on light absorption in mixed canopies, Ann. For. Sci., 70(2), 123–131, doi:10.1007/s13595-012-0246-8, 2013.

Šimůnek, J., Šejna, M., Saito, H., Sakai, M. and van Genuchten, M. T.: The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media, Version 4.08, University of California Riverside, Riverside, USA., 2009.

Sitch, S., Smith, B., Prentice, I. C., Arneth, a., Bondeau, a., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K. and Venevsky, S.: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model , Glob. Chang. Biol., 9(2), 161–185, doi:10.1046/j.1365-2486.2003.00569.x, 2003.

Smith, K. A.: Investigating Uncertainty in Global Hydrology Modelling, Ph.D.thesis, University of Nottingham, Nottingham, UK., 2016.

Sperna Weiland, F. C., Vrugt, J. A., Van Beek, R. L. P. H., Weerts, A. H. and Bierkens, M. F. P.: Significant uncertainty in global scale hydrological modeling from precipitation data errors, J. Hydrol., 529(3), 1095–1115, doi:10.1016/j.jhydrol.2015.08.061, 2015.

Stewart, J. B.: Modelling surface conductance of pine forest, Agric. For. Meteorol., 43(1), 19–35, doi:10.1016/0168-1923(88)90003-2, 1988.

Sutanudjaja, E. H., Van Beek, L. P. H., De Jong, S. M., Van Geer, F. C. and Bierkens, M. F. P.: Large-scale groundwater modeling using global datasets: A test case for the Rhine-Meuse basin, Hydrol. Earth Syst. Sci., 15(9), 2913–2935, doi:10.5194/hess-15-2913-2011, 2011.

Taylor, R. G., Todd, M. C., Kongola, L., Maurice, L., Nahozya, E., Sanga, H. and MacDonald, A. M.: Evidence of the dependence of groundwater resources on extreme rainfall in East Africa, Nat. Clim. Chang., 3, 374–378, doi:10.1038/nclimate1731, 2013.

Tesemma, Z. K., Wei, Y., Peel, M. C. and Western, A. W.: The effect of year-to-year variability of leaf area index on Variable Infiltration Capacity model performance and simulation of runoff, Adv. Water Resour., 83, 310–322, doi:10.1016/j.advwatres.2015.07.002, 2015.

Thornthwaite, C. W.: An Approach toward a Rational Classification of Climate, Geogr. Rev., 38(1), 55–94, doi:10.2307/210739, 1948.

Tritz, S., Guinot, V. and Jourde, H.: Modelling the behaviour of a karst system catchment using non-linear hysteretic conceptual model, J. Hydrol., 397(3–4), 250–262, doi:10.1016/j.jhydrol.2010.12.001, 2011.

Twine, T. E., Kustas, W. P., Norman, J. M., Cook, D. R., Houser, P. R., Meyers, T. P., Prueger, J. H., Starks, P. J. and Wesley, M. L.: Correcting eddy covariance flux underestimates over grassland, Agric. For. Meteorol., 103(3), 279–300, doi:10.1016/S0168-1923(00)00123-4, 2000.

Uhlenbrook, S.: Catchment hydrology—a science in which all processes are preferential, Hydrol. Process., 20(16), 3581–3585, doi:10.1002/hyp.6564, 2006.

Valente, F., David, J. S. and Gash, J. H. C.: Modelling interception loss for two sparse eucalypt and pine forests in central Portugal using reformulated Rutter and Gash analytical models, J. Hydrol., 190(1–2), 141–162, doi:10.1016/S0022-1694(96)03066-1, 1997.

Van Beek, L. P. H. and Bierkens, M. F. P.: The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification, Report Department of Physical Geography, Utrecht University, Netherlands. [online] Available from: http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf, 2008.

Van Beek, R.: Forcing PCR-GLOBWB with CRU data, Utrecht University, the Netherlands. [online] Available from: http://vanbeek.geo.uu.nl/suppinfo/vanbeek2008.pdf, 2008.

Van Dijk, A. I. J. M. and Bruijnzeel, L. A.: Modelling rainfall interception by vegetation of variable density using an adapted analytical model. Part 1. Model description, J. Hydrol., 247(3–4), 230–238, doi:10.1016/S0022-1694(01)00392-4, 2001.

Van Werkhoven, K., Wagener, T., Reed, P. and Tang, Y.: Rainfall characteristics define the value of streamflow observations for distributed watershed model identification, Geophys. Res. Lett., 35(11), L11403, doi:10.1029/2008GL034162, 2008.

Vörösmarty, C. J.: Global change , the water cycle, and our search for Mauna Loa, Hydrol. Process., 16(1), 135–139, doi:10.1002/hyp.527, 2002.

Vörösmarty, C. J., Moore, B., Grace, A. L., Gildea, M. P., Melillo, J. M., Peterson, B. J., Rastetter, E. B. and Steudler, P. A.: Continental scale models of water balance and fluvial transport: An application to South America, Global Biogeochem. Cycles, 3(3), 241–265, doi:10.1029/GB003i003p00241, 1989.

Vörösmarty, C. J., Federer, C. A. and Schloss, A. L.: Potential evapotranspiration functions compared on US watersheds: implications for global-scale water balance and terrestrial ecosystem modeling, J. Hydrol., 207(3–4), 147–169, doi:10.1016/S0022-1694(98)00109-7, 1998.

Wada, Y., Van Beek, L. P. H. and Bierkens, M. F. P.: Nonsustainable groundwater sustaining irrigation: A global assessment, Water Resour. Res., 48(6), W00L06, doi:10.1029/2011WR010562, 2012.

Wang, K. and Dickinson, R. E.: A review of global terrestrial evapotranspiration: observation, modelling, climatology, and climatic variability, Rev. Geophys., 50(2), 1–54, doi:10.1029/2011RG000373, 2012.

Weiler, M. and McDonnell, J.: Virtual experiments: A new approach for improving process conceptualization in hillslope hydrology, J. Hydrol., 285(1–4), 3–18, doi:10.1016/S0022-1694(03)00271-3, 2004.

Werth, S., Güntner, A., Petrovic, S. and Schmidt, R.: Integration of GRACE mass variations into a global hydrological model, Earth Planet. Sci. Lett., 277(1–2), 166–173, doi:10.1016/j.epsl.2008.10.021, 2009.

Williams, P. W.: The role of the subcutaneous zone in karst hydrology, J. Hydrol., 61(1–3), 45–67, doi:10.1016/0022-1694(83)90234-2, 1983.

Williams, P. W.: Environmental change and human impact on karst terrains: an introduction, in Karst Terrains - Environmental changes and human impact, edited by P. W. Williams, Catena Supplement 25, Catena Verlag, Cremlingen-Destedt, Germany., 1–19, 1993.

Williams, P. W.: The role of the epikarst in karst and cave hydrogeology : a review, Int. J. Speleol., 37, 1–10, doi:10.5038/1827-806X.37.1.1, 2008.

Yin, J., Porporato, A. and Albertson, J.: Interplay of climate seasonality and soil moisture-rainfall feedback, Water Resour. Res., 50(7), 6053–6066, doi:10.1002/2013WR014772, 2014.

Young, P., Parkinson, S. and Lees, M.: Simplicity out of complexity in environmental modelling: Occam's razor revisited, J. Appl. Stat., 23(2–3), 165–210, doi:10.1080/02664769624206, 1996.

Zhang, Z., Chen, X., Ghadouani, A. and Shi, P.: Modelling hydrological processes influenced by soil, rock and vegetation in a small karst basin of southwest China, Hydrol. Process., 25(15), 2456–2470, doi:10.1002/hyp.8022, 2011.

Zhu, Z., Piao, S., Myneni, R. B., Huang, M., Zeng, Z., Canadell, J. G., Ciais, P., Sitch, S., Friedlingstein, P., Arneth, A., Cao, C., Cheng, L., Kato, E., Koven, C., Li, Y., Lian, X., Liu, Y., Liu, R., Mao, J., Pan, Y., Peng, S., Peñuelas, J., Poulter, B., Pugh, T. A. M., Stocker, B. D., Viovy, N., Wang, X., Wang, Y., Xiao, Z., Yang, H., Zaehle, S. and Zeng, N.: Greening of the Earth and its drivers, Nat. Clim. Chang., 6, 791–795, doi:10.1038/nclimate3004, 2016.

Parameter	Description	unit	Lower limit	Upper limit	Category
h_{veg}	Vegetation height	[m]	0.2	Site specific	vegetation
r_{st}	Stomatal resistance	[s m^{-1}]	20	600	vegetation
LAI_{min}	Reduction in leaf area index during the dormant season	[%]	5	100	vegetation
LAI_{max}	Annual maximum leaf area index	[m^{2} m^{-2}]	0.5	8	vegetation
V_r	Maximum storage capacity of the root zone	[mm]	20	500	vegetation
V_{can}	Canopy storage capacity per unit of LAI	[mm LAI]	0.1	0.5	vegetation
k	Beer-Lambert's law extinction coefficient	[-]	0.4	0.7	vegetation
f_{red}	Reduction factor for transpiration below the root zone	[-]	0	0.15	soil
z_0	Soil roughness length	[m]	0.0003	0.013	soil
$r_{s,soi}$	Soil surface resistance	[s m^{-1}]	0	100	soil
V_e	Maximum storage capacity of the first soil layer	[mm]	5	45	soil
a	Spatial variability coefficient	[-]	0	6	soil and epikarst
V_{soil}	Mean soil storage capacity	[mm]	20	800	soil
V_{epi}	Mean epikarst storage capacity	[mm]	200	700	epikarst
K_{epi}	Mean epikarst outflow coefficient	[d]	0	50	epikarst

Table 1. Description of V2Karst parameters, unconstrained ranges used in the application at the four FLUXNET sites to capture the variability across soil, epikarst and vegetation types, category of the parameters (which indicated whether the parameters depend on soil, epikarst or vegetation properties). Parameters a, V_{soil}, V_{epi} and K_{epi} were already present in the previous version of the model (VarKarst). More information on how the ranges were determined is provided in Sect. S3 of our supplementary material.

Site	Simulation period (including a one-year warm-up period)		Number of months with latent heat measurement for calibration	Number of months with soil moisture measurement for calibration
	Start	End		
German site	1 Jan. 2000	17 Dec. 2009	62	74
Spanish site	1 Jan. 2005	30 Dec. 2011	12	12
French 1 site	2 Jan. 2009	30 Dec. 2011	13	Not measured
French 2 site	18 Apr. 2002	29 Jun. 2009	37	Not measured

Table 2. Simulation period at the four FLUXNET sites, and number of months where latent heat measurements and soil moisture measurements are available to calibrate the model. Soil moisture measurements are not provided at the two French sites.

Parameter	Unit	German site (deciduous forest)		Spanish site (shrubland)		French 1 site (evergreen forest)		French 2 site (evergreen forest)	
		Lower limit	Upper limit	Lower limit	Upper limit	Lower limit	Upper limit	Lower limit	Upper limit
h_{veg}	[m]	23.1	42.9	0.35	0.85	7.1	13.3	3.9	7.2
r_{st}	[s m^{-1}]	275	400	195	350	320	455	320	455
LAI_{min}	[%]	5	20	34	63	80	100	80	100
LAI_{max}	[m^2 m^{-2}]	3.5	6.5	1.9	3.5	1.5	2.9	2.0	3.8
V_r	[mm]	60	300	30	200	30	200	30	200
V_{soi}	[mm]	60	400	30	300	30	300	30	300

Table 3. Site-specific constrained parameter ranges at the four FLUXNET sites for the vegetation parameters (h_{veg}, r_{st}, LAI_{min}, LAI_{max}, V_r) and for the soil storage capacity (V_{soi}). More information on how the ranges were determined is provided in Sect. S3 of our supplementary material. Parameters are defined in Table 1.

V2Karst input		Unit	Virtual forest site	Virtual shrub site
Vegetation parameter	h_{veg}	[m]	32.1	0.4
	r_{st}	[s m^{-1}]	390	291
	LAI_{min}	[%]	16	38
	LAI_{max}	[m^{2} m^{-2}]	5.0	2.0
	V_r	[mm]	289	151
	V_{can}	[mm LAI]	0.29	0.35
	k	[-]	0.53	0.45
Soil and epikarst parameter	f_{red}	[-]	0.010	0.080
	z_0	[m]	0.0110	0.0045
	$r_{s,soi}$	[s m^{-1}]	56	61
	V_e	[mm]	11	8
	a	[-]	1.8	1.9
	V_{soil}	[mm]	373	174
	V_{epi}	[mm]	396	519
	K_{epi}	[d]	33	15
Weather input (winter)	R_n	[MJ m^{-2} d^{-1}]	-0.0	2.2
	T	[°C]	0.1	4.9
	RH	[%]	89	61
	WS [a]	[m s^{-1}]	3.5	4.0
Weather input (summer)	R_n	[MJ m^{-2} d^{-1}]	10.5	12.1
	T	[°C]	16.6	20.4
	RH	[%]	72	43
	WS [a]	[m s^{-1}]	2.6	3.4

Table 4. Values of V2Karst parameters and weather variables used in the virtual experiment. Values for the virtual forest site and the virtual shrub site are based on the characteristics of the German FLUXNET site and Spanish FLUXNET site respectively. Values of the model parameters (parameters are defined in Table 1) correspond to behavioural parameterisations obtained when calibrating the model and values of the weather variables (R_n net radiation, T temperature, RH relative humidity, WS wind speed) correspond to the average values calculated at FLUXNET sites.

[a] At the virtual shrub site, WS was recalculated at a height of 43.5 m because the original measurement provided at a height of 2.5 m at the Spanish site was too low to simulate a change of land cover to tall vegetation (forest). More details on this are reported in Section S4 of our supplementary material.

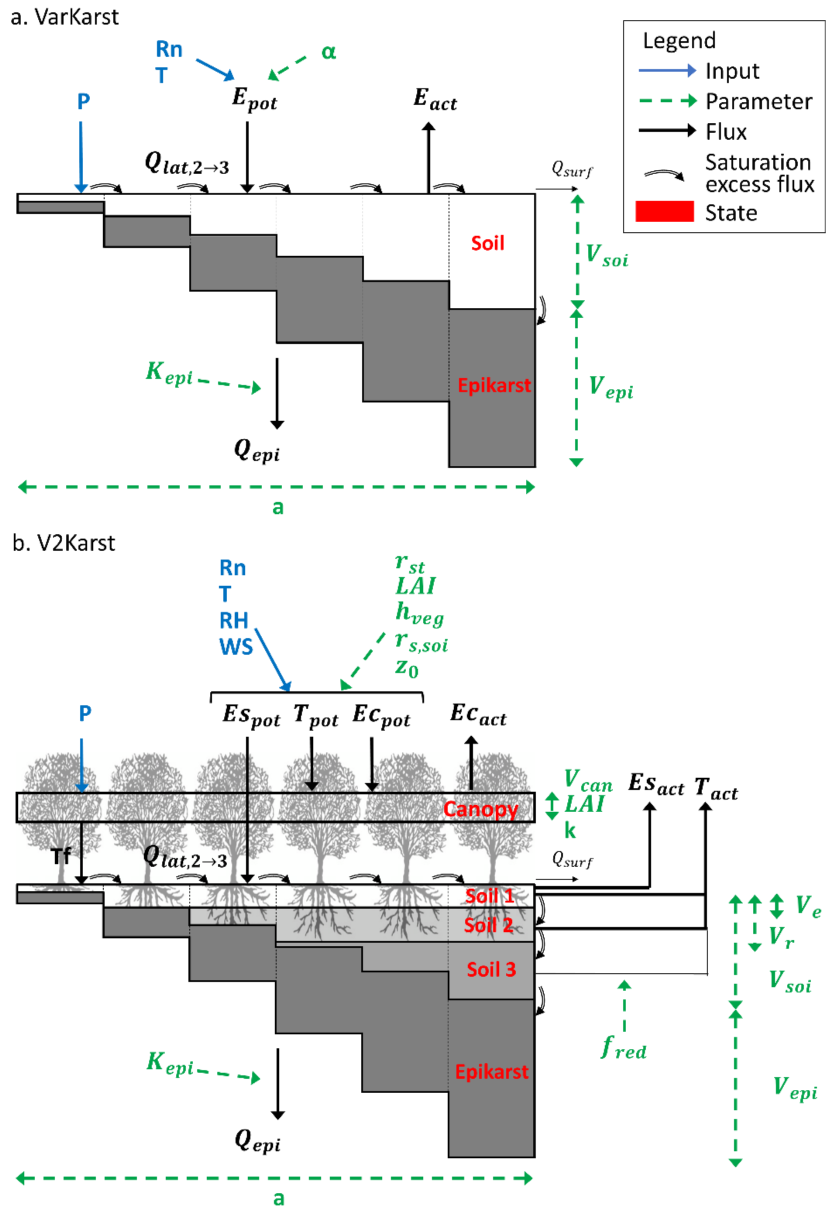


Figure 1. Schematic representation of (a) the VarKarst model (Hartmann et al., 2015) and (b) the new version of the model V2Karst using six vertical compartments. Model parameters are in green (see their definition in Table 1**)**, inputs are in blue (P precipitation, R_n net radiation, T temperature, RH relative humidity, WS wind speed), model fluxes are in black (E_{pot} potential total evapotranspiration, T_{pot} potential transpiration, Ec_{pot} potential evaporation from canopy interception, Es_{pot} potential soil evaporation, E_{act} total actual ET, T_{act} actual transpiration, Ec_{act} actual evaporation from canopy interception, Es_{act} actual bare soil evaporation, Tf throughfall, $Q_{lat,2\to3}$ lateral flow from the second to the third compartment, Q_{surf} surface runoff and Q_{epi} recharge) and state variables are in red.

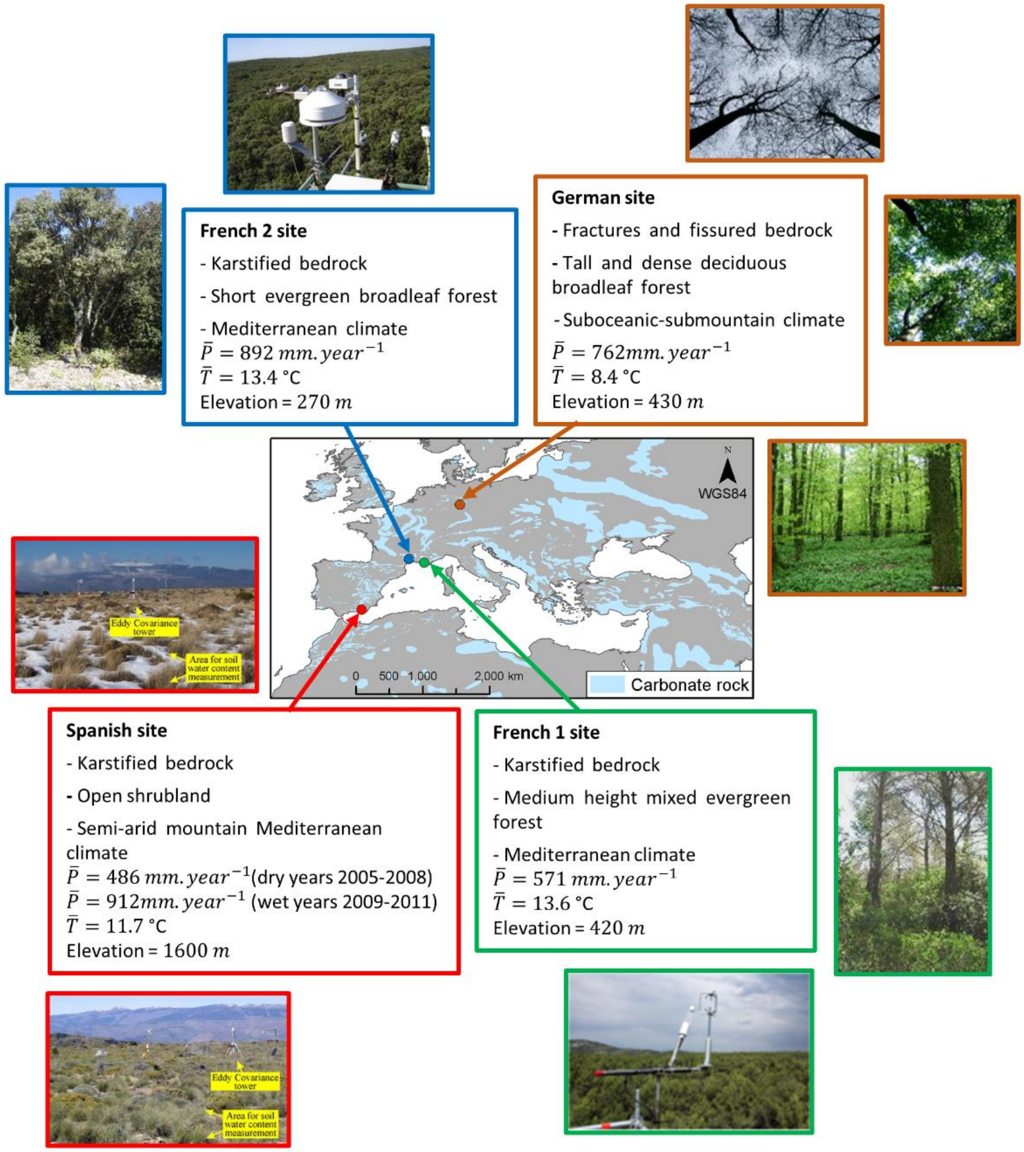


Figure 2. Four carbonate rock FLUXNET sites selected for the analyses. Mean annual precipitation $\bar{P}$ and mean annual temperature $\bar{T}$ were estimated over the period 1 January 2001-17 December 2009 for the German site, 1 January 2006-31 December 2008 for the Spanish site (dry years), 1 January 2009-30 December 2011 for the Spanish site (wet years), 1 January 2010-30 December 2011 for the French 1 site and 1 April 2003-31 March 2009 for the French 2 site.

Sources of the photos: (Pinty et al., 2011) for the German site, (Alcalá et al., 2011) for the Spanish site, http://www.gip-ecofor.org/f-ore-t/fontBlanche.php for the French 1 site, http://puechabon.cefe.cnrs.fr/ for the French 2 site. Source of the carbonate rock and country map: (Williams and Ford, 2006) (country map obtained from Terraspace, Russian space agency).

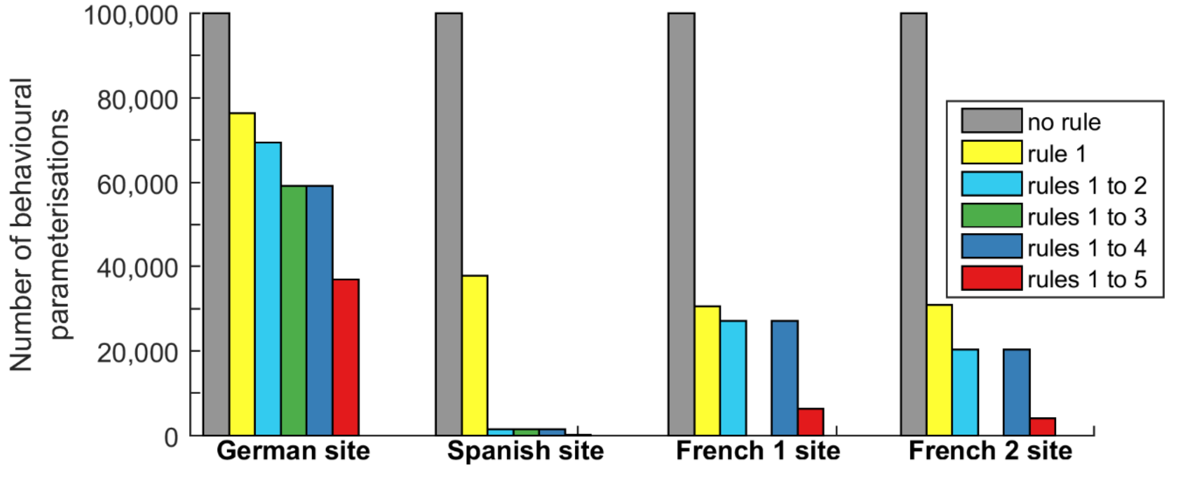


Figure 3. Reduction in the number of behavioural parameterisations of the V2Karst model at the four FLUXNET sites, when applying sequentially the five soft rules defined in Sect. 4.1 (no rule: initial sample; rule 1: ET bias; rule 2: ET correlation; rule 3: soil moisture correlation; rule 4: runoff; rule 5: a priori information). Rule 3 could not be applied to the French sites where soil moisture observations are not available.

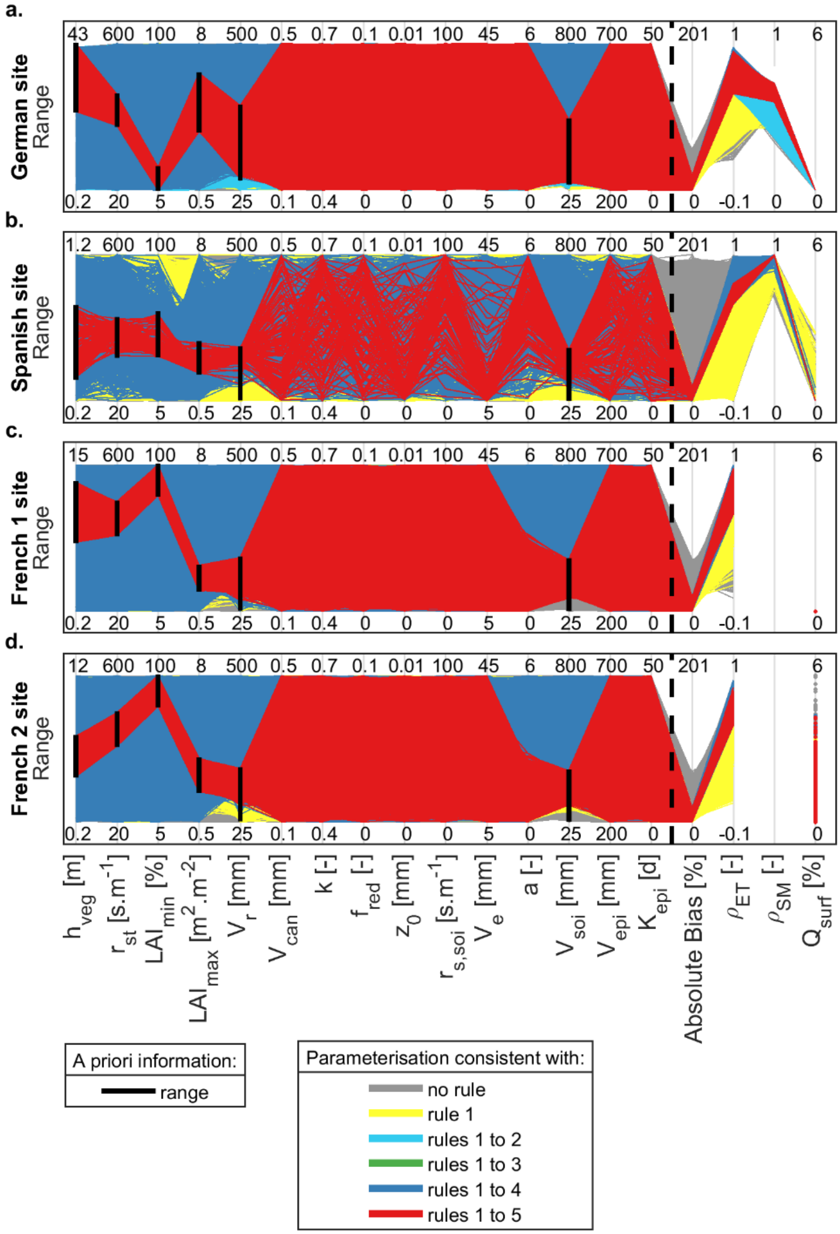


Figure 4. Parallel coordinate plots representing V2Karst behavioural parameterisations, and their corresponding simulated output values, identified when sequentially applying the five soft rules defined in Sect. 4.1 at (a) the German site, (b) the Spanish site, (c) the French 1 site and (d) the French 2 site. Parameters are defined in Table 1. $BIAS$ absolute mean error between observed and simulated total actual ET (rule 1), ρ_{ET} correlation coefficient between observed and simulated total actual ET (rule 2), ρ_{SM} correlation coefficient between observed and simulated soil moisture (rule 3), Q_{surf} surface runoff (rule 4). Rule 5 corresponds to application of a priori information on parameter ranges (black vertical bars).

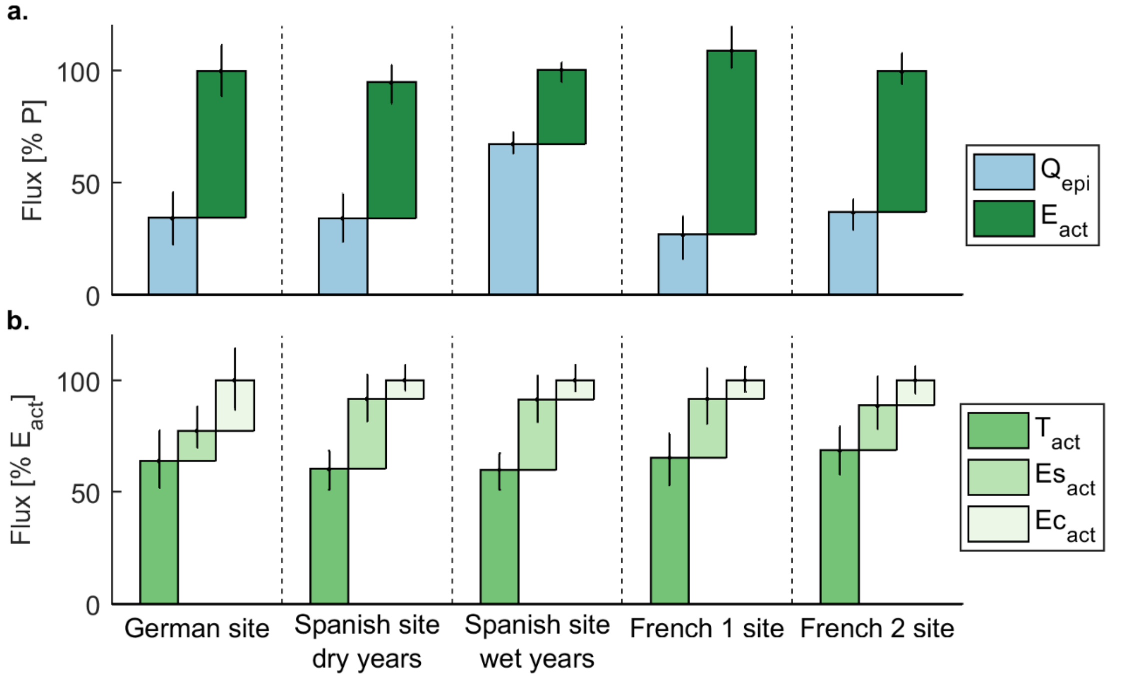
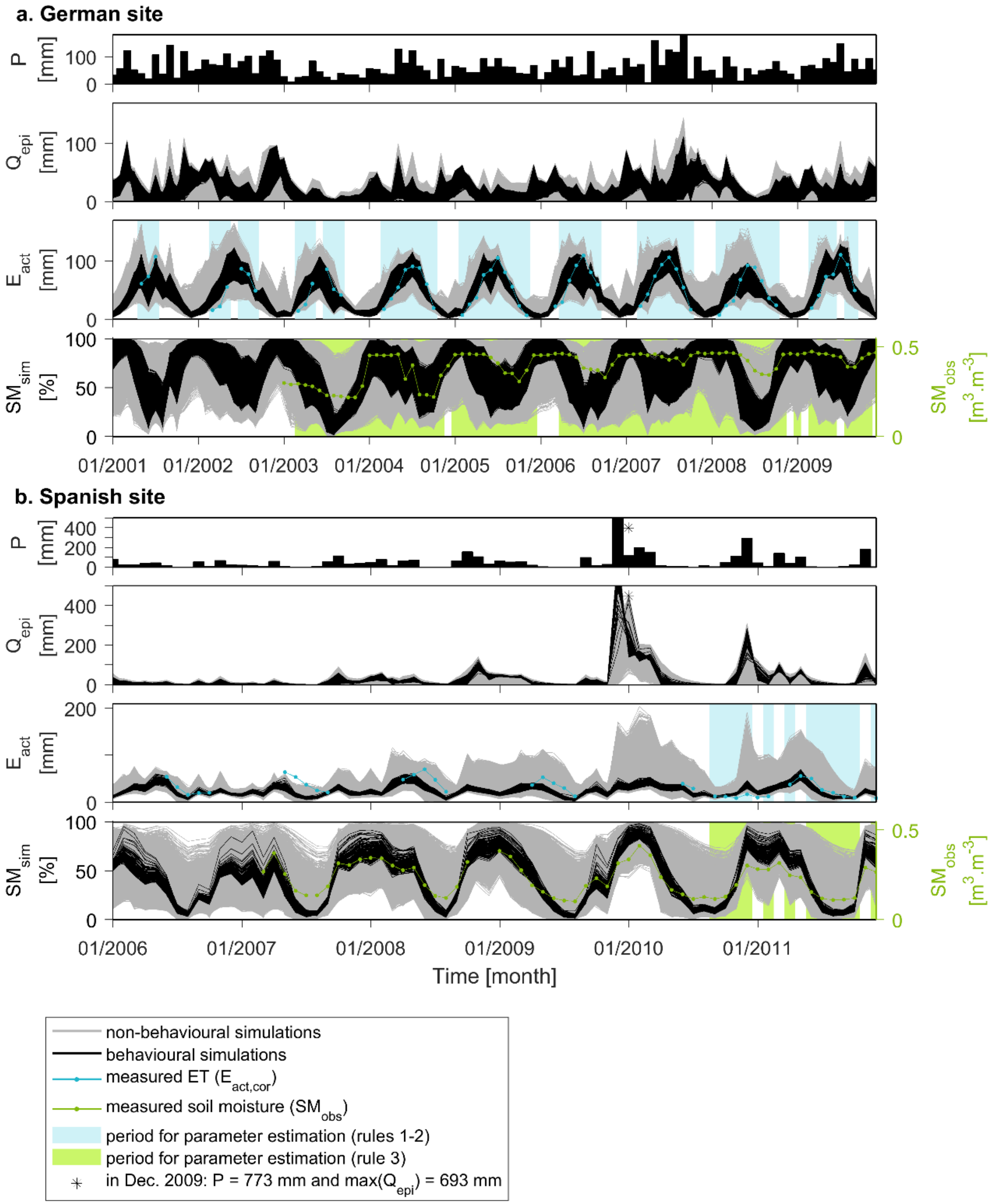


Figure 5. (a) Simulated recharge (Q_{epi}) and actual ET (E_{act}) expressed as a percentage of total precipitation and (b) simulated actual transpiration (T_{act}), actual soil evaporation (Es_{act}) and actual evaporation from interception (Ec_{act}) expressed as a percentage of E_{act}. The figure reports the ensemble mean and 95% confidence intervals calculated over the behavioural simulation ensemble of the V2Karst model at the four FLUXNET sites. Simulated fluxes were evaluated over the period 1 January 2001-17 December 2009 for the German site, 1 January 2006-31 December 2008 for the Spanish site (dry years), 1 January 2009-30 December 2011 for the Spanish site (wet years), 1 January 2010-30 December 2011 for the French 1 site and 1 April 2003-31 March 2009 for the French 2 site.



a. German site
P [mm]
100
0
Qepi [mm]
100
0
Eact [mm]
100
0
SMsim [%]
100
50
0
SMobs [m3.m-3]
0.5
0
01/2001 01/2002 01/2003 01/2004 01/2005 01/2006 01/2007 01/2008 01/2009
b. Spanish site
P [mm]
400
200
0
Qepi [mm]
400
200
0
Eact [mm]
200
0
SMsim [%]
100
50
0
SMobs [m3.m-3]
0.5
0
01/2006 01/2007 01/2008 01/2009 01/2010 01/2011
Time [month]
non-behavioural simulations
behavioural simulations
measured ET (Eact,cor)
measured soil moisture (SMobs)
period for parameter estimation (rules 1-2)
period for parameter estimation (rule 3)
in Dec. 2009: P = 773 mm and max(Qepi) = 693 mm

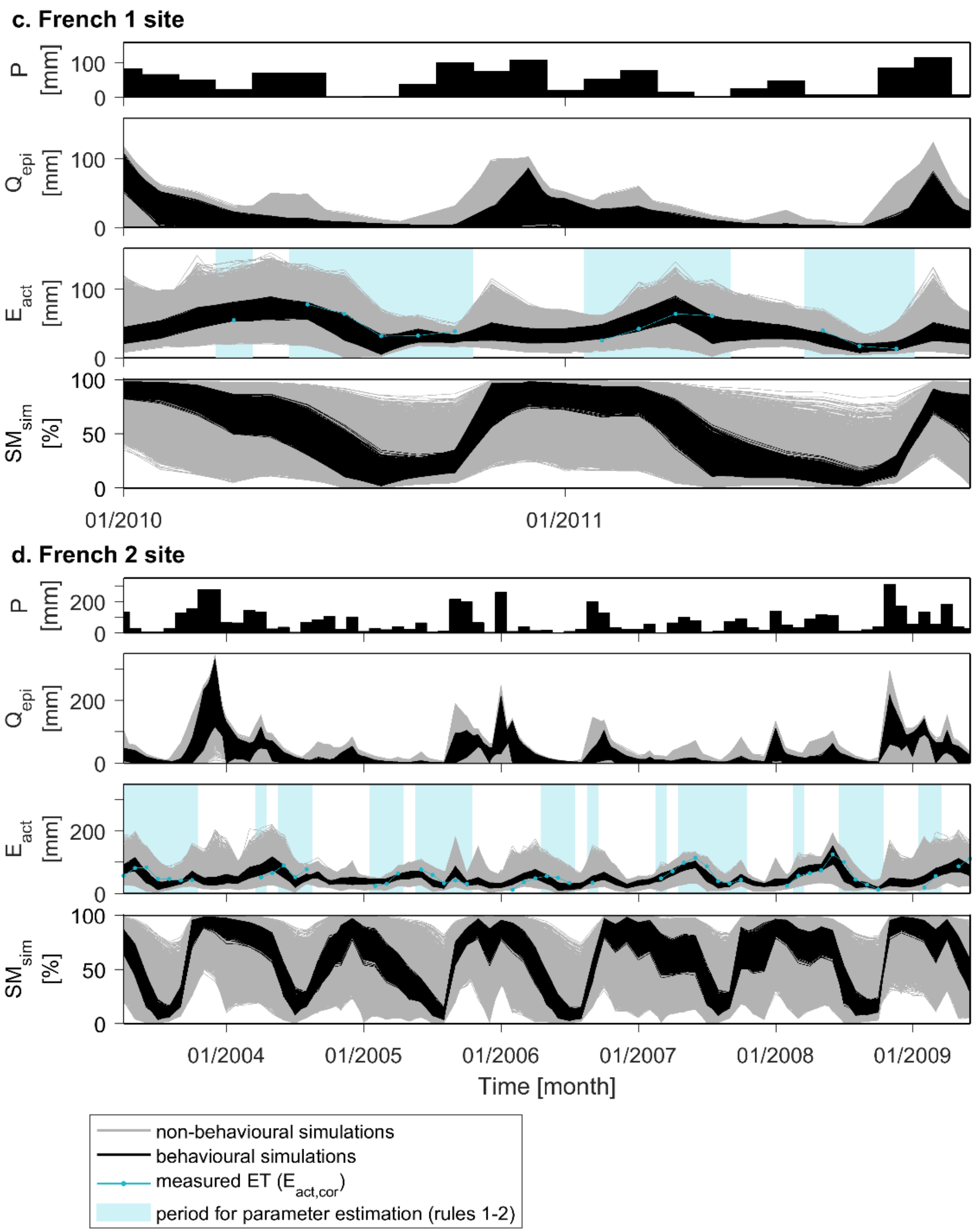


Figure 6. Monthly time series of observed variables, namely precipitation input P, actual $E_{act,cor}$ (in blue) and soil moisture SM_{obs} (in green), and monthly time series of simulated variables namely recharge Q_{epi}, actual ET $E_{act,sim}$ and soil moisture in the root zone SM_{sim} (non-behavioural simulations are in grey and behavioural simulations in black)~~precipitation input (P), simulated recharge (Q_{epi}), simulated actual ET (E_{act}, which is the sum of evaporation from canopy interception, transpiration and soil evaporation), simulated soil moisture within the root zone (SM_{sim}), and monthly observations of actual ET and soil moisture~~ at (a) the German site, (b) the Spanish site, (c) the French 1 site and (d) the French 2 site. Blue and ~~red~~ green shaded

areas correspond to the periods in which observation of ET and soil moisture respectively were selected to apply the soft rules of Sect. 4.1 (further details on data processing in Sect. 3.2).

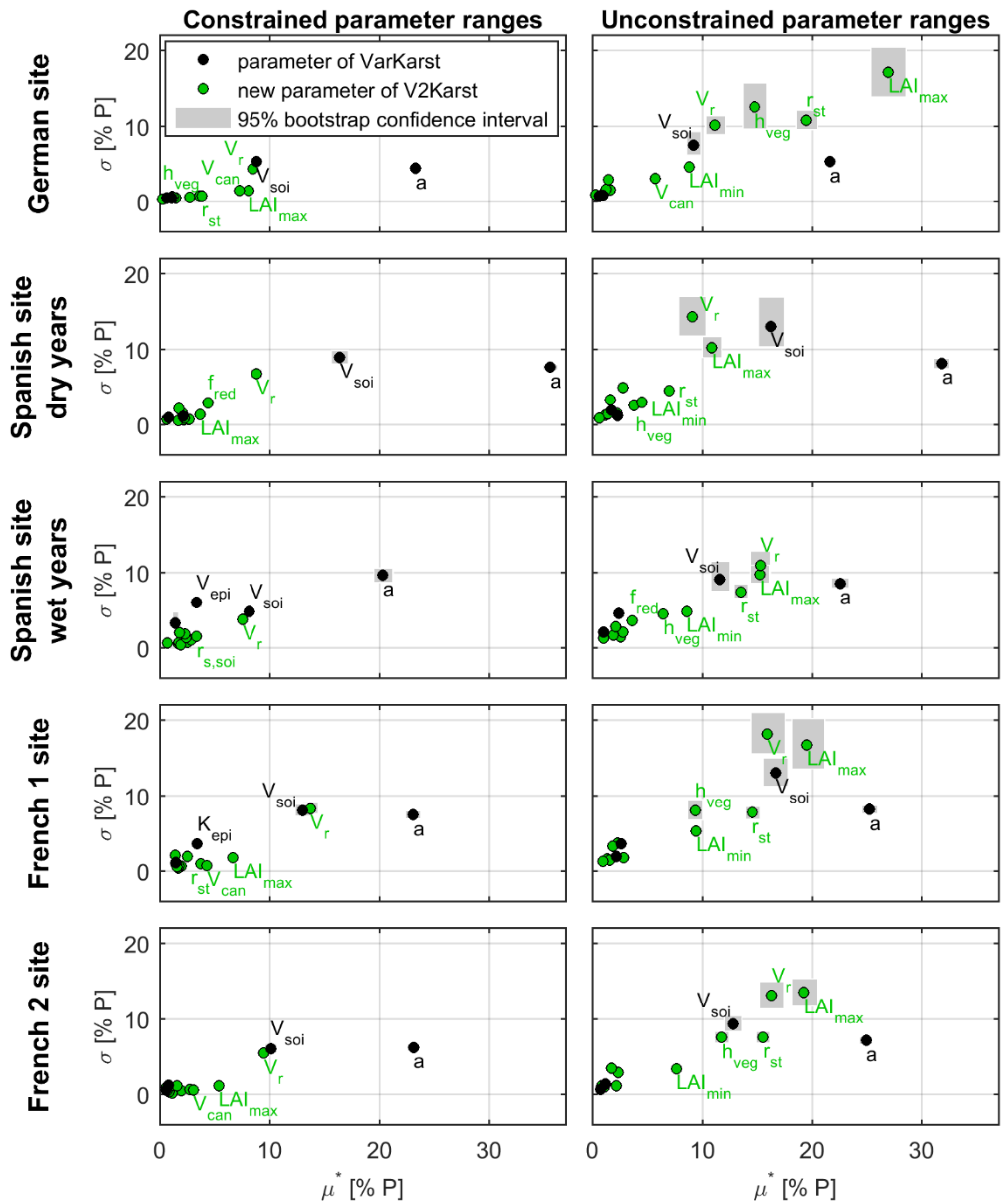


Figure 7. Sensitivity indices of the V2Karst parameters (μ^* is the mean of the absolute Elementary Effects and σ is the standard deviation of the Elementary Effects) for total simulated recharge (expressed as a percentage of total precipitation) at the four FLUXNET sites, when constrained (site-specific) parameter ranges are used (ranges of Table 3) and when

unconstrained ranges are used (ranges of Table 1). Sensitivity indices were computed over the period 1 January 2001-17 December 2009 for the German site, 1 January 2006-31 December 2008 for the Spanish site (dry years), 1 January 2009-30 December 2011 for the Spanish site (wet years), 1 January 2010-30 December 2011 for the French 1 site and 1 April 2003-31 March 2009 for the French 2 site.

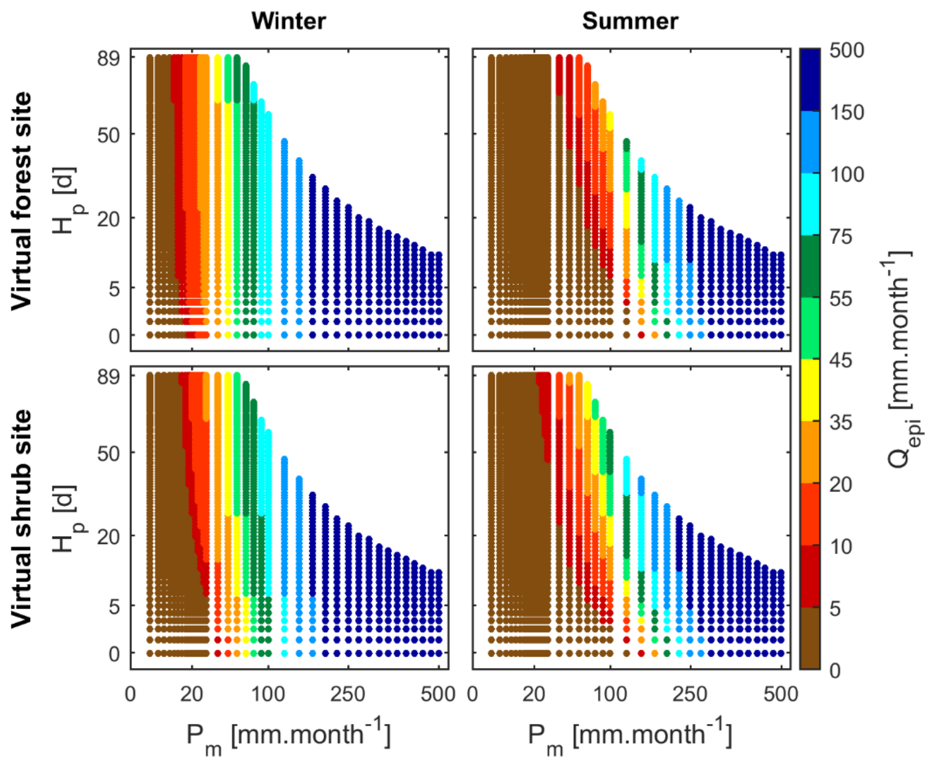


Figure 8. Average monthly recharge (Q_{epi}) simulated with V2Karst for different values of the average monthly precipitation amount P_m [mm month^{-1}] and the interval between wet days H_p [d] of the synthetic periodic precipitation input used to force the model at the virtual forest and shurb sites and under winter and summer conditions.

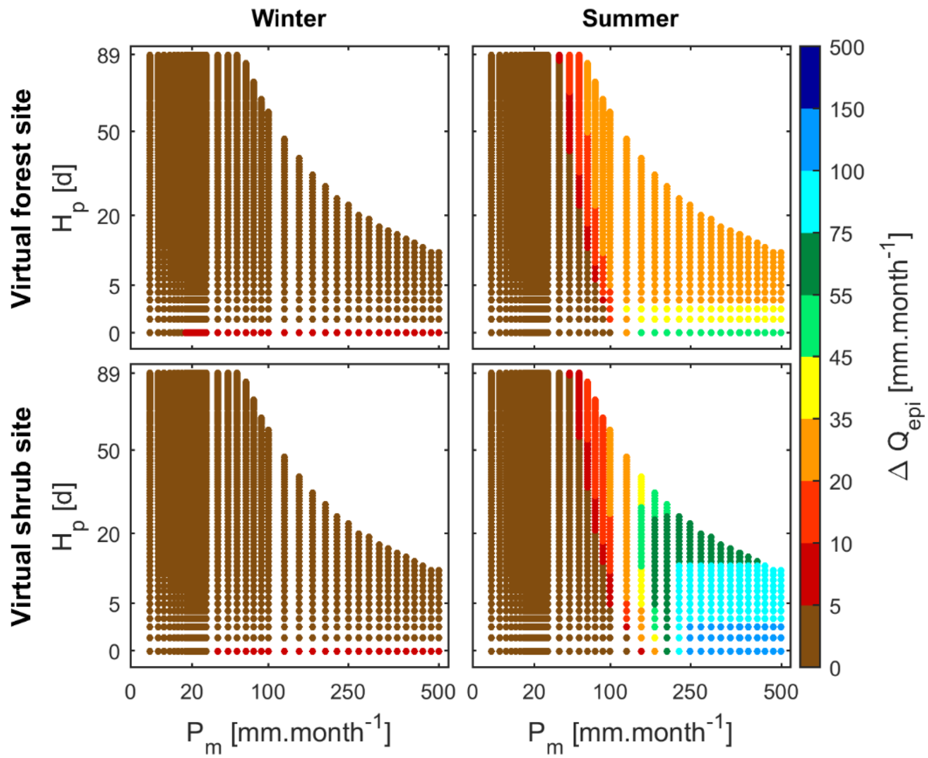


Figure 9. Change in monthly recharge ($\Delta Q_{epi} = Q_{epi}^{shrub} - Q_{epi}^{forest}$) simulated with V2Karst when the land cover is set to shrub compared to forest for different values of the average monthly precipitation amount P_m [mm ~~.~~month^{-1}] and the interval between wet days ~~(~~H_p [d] of the synthetic periodic precipitation input used to force the model at the virtual forest and shurb sites and under winter and summer conditions.